

Table 1: Cleavage of 75 human light chains.

	Enzyme	Recognition*	Nch	Ns	Planned location of site
	AfeI	AGCgct	0	0	
5	AflII	Cttaag	0	0	HC FR3
	AgeI	Accggt	0	0	
	AscI	GGGcgccc	0	0	After LC
	BglII	Agatct	0	0	
	BsiWI	Cgtacg	0	0	
	BspDI	ATcgat	0	0	
10	BssHII	Gcgcg	0	0	
	BstBI	TTcgaa	0	0	
	DraIII	CACNNNgtg	0	0	
	EagI	Cggccg	0	0	
	FseI	GGCCGGcc	0	0	
15	FspI	TGCgca	0	0	
	HpaI	GTTaac	0	0	
	MfeI	Caattg	0	0	HC FR1
	MluI	Acgctg	0	0	
	NcoI	Ccatgg	0	0	Heavy chain signal
20	NheI	Gctagc	0	0	HC/anchor linker
	NotI	GCggccgc	0	0	In linker after HC
	NruI	TCGcga	0	0	
	PacI	TTAATtaa	0	0	
	PmeI	GTTTaaac	0	0	
25	PmlI	CACgtg	0	0	
	PvuI	CGATcg	0	0	
	SacII	CCGCgg	0	0	
	SalI	Gtcgac	0	0	
	SfiI	GGCCNNNNnggcc	0	0	Heavy Chain signal
30	SgfI	GCGATcgc	0	0	
	SnaBI	TACgta	0	0	
	StuI	AGGcct	0	0	
	XbaI	Tctaga	0	0	HC FR3
	AatII	GACGTc	1	1	
35	AclI	AACgtt	1	1	
	AseI	ATtaat	1	1	
	BsmI	GAATGCN	1	1	
	BspEI	Tccgga	1	1	HC FR1
	BstXI	CCANNNNntgg	1	1	HC FR2
40	DrdI	GACNNNNngtc	1	1	
	HindIII	Aagctt	1	1	
	PciI	Acatgt	1	1	
	SapI	gaagagc	1	1	
	ScaI	AGTact	1	1	
45	SexAI	Accwgg	1	1	
	SpeI	Actagt	1	1	
	TliI	Ctcgag	1	1	
	XhoI	Ctcgag	1	1	
	BcgI	cgannnnntgc	2	2	
50	BlpI	GCtnagc	2	2	
	BssSI	Ctcgtg	2	2	
	BstAPI	GCANNNNntgc	2	2	
	EspI	GCtnagc	2	2	
	KasI	Ggcgcc	2	2	
55	PflMI	CCANNNNntgg	2	2	
	XmnI	GAANNnttc	2	2	

	<b>ApaI</b>	<b>Gtgcac</b>	<b>3</b>	<b>3</b>	<b>LC signal seq</b>
	NaeI	GCCggc	3	3	
	NgoMI	Gccggc	3	3	
5	PvuII	CAGctg	3	3	
	RsrII	CGGwccg	3	3	
	BsrBI	GAGcgg	4	4	
	BsrDI	GCAATGNNn	4	4	
	BstZ17I	GTAtac	4	4	
10	EcoRI	Gaattc	4	4	
	SphI	GCATGc	4	4	
	SspI	AATatt	4	4	
	AccI	GTMkac	5	5	
	BclI	Tgatca	5	5	
	BsmBI	Nnnnnngagacg	5	5	
15	BsrGI	Tgtaca	5	5	
	DraI	TTTaaa	6	6	
	<b>NdeI</b>	<b>CATatg</b>	<b>6</b>	<b>6</b>	<b>HC FR4</b>
	SwaI	ATTTaaat	6	6	
20	BamHI	Ggatcc	7	7	
	SacI	GAGCTc	7	7	
	BciVI	GTATCCNNNNNN	8	8	
	BsaBI	GATNNnnatc	8	8	
	NsiI	ATGCAt	8	8	
	<b>Bsp120I</b>	<b>Gggccc</b>	<b>9</b>	<b>9</b>	<b>CH1</b>
25	<b>ApaI</b>	<b>GGGCCc</b>	<b>9</b>	<b>9</b>	<b>CH1</b>
	PspOoMI	Gggccc	9	9	
	BspHI	Tcatga	9	11	
	EcoRV	GATatc	9	9	
	AhdI	GACNNNnngtc	11	11	
30	BbsI	GAAGAC	11	14	
	PsiI	TTATAa	12	12	
	BsaI	GGTCTCNnnnn	13	15	
	XmaI	Cccggg	13	14	
	AvaI	Cycgrg	14	16	
35	BglI	GCCNNNNnnggc	14	17	
	AlwNI	CAGNNNctg	16	16	
	BspMI	ACCTGC	17	19	
	XcmI	CCANNNNNnnntgg	17	26	
40	<b>BstEII</b>	<b>Ggtnacc</b>	<b>19</b>	<b>22</b>	<b>HC FR4</b>
	Sse8387I	CCTGCAGg	20	20	
	AvrII	Cctagg	22	22	
	HincII	GTYrac	22	22	
	BsgI	GTGCAG	27	29	
	MscI	TGGcca	30	34	
45	BseRI	NNnnnnnnnnctcctc	32	35	
	Bsu36I	CCtnagg	35	37	
	PstI	CTGCAG	35	40	
	EciI	nnnnnnnnntccgcc	38	40	
	PpuMI	RGgwccy	41	50	
50	StyI	Ccwwgg	44	73	
	EcoO109I	RGgnccy	46	70	
	Acc65I	Ggtacc	50	51	
	KpnI	GGTACc	50	51	
	BpmI	ctccag	53	82	
55	AvaII	Ggwcc	71	124	

\* cleavage occurs in the top strand after the last upper-case base. For REs

that cut palindromic sequences, the lower strand is cut at the symmetrical site.

Table 2: Cleavage of 79 human heavy chains

	Enzyme	Recognition	Nch	Ns	Planned location of site
5	AfeI	AGCgct	0	0	
	AflII	Cttaag	0	0	HC FR3
	AscI	GGGcgccc	0	0	After LC
10	BsiWI	Cgtacg	0	0	
	BspDI	ATcgat	0	0	
	BssHII	Gcggcg	0	0	
	FseI	GGCCGGcc	0	0	
	HpaI	GTTaac	0	0	
15	NheI	Gctagc	0	0	HC Linker
	NotI	GCggccgc	0	0	In linker, HC/anchor
	NruI	TCGcga	0	0	
	NsiI	ATGCAt	0	0	
	PacI	TTAATtaa	0	0	
20	PciI	Acattg	0	0	
	PmeI	GTTTaaac	0	0	
	PvuI	CGATcg	0	0	
	RsrII	CGgwcgc	0	0	
	SapI	gaagagc	0	0	
25	SfiI	GGCCNNNNnggcc	0	0	HC signal seq
	SgfI	GCGATcgc	0	0	
	SwaI	ATTTaaat	0	0	
	AclI	AAcggt	1	1	
	AgeI	Accggt	1	1	
30	AseI	ATtaat	1	1	
	AvrII	Cctagg	1	1	
	BsmI	GAATGCN	1	1	
	BsrBI	GAGcgg	1	1	
	BsrDI	GCAATGNNn	1	1	
35	DraI	TTTaaa	1	1	
	FspI	TGCgca	1	1	
	HindIII	Aagctt	1	1	
	MfeI	Caattg	1	1	HC FR1
	NaeI	GCCggc	1	1	
40	NgoMI	Gccggc	1	1	
	SpeI	Actagt	1	1	
	Acc65I	Ggtacc	2	2	
	BstBI	TTcgaa	2	2	
45	KpnI	GGTACc	2	2	
	MluI	Acgcgt	2	2	
	NcoI	Ccatgg	2	2	In HC signal seq
	NdeI	CAtatg	2	2	HC FR4
	PmlI	CACgtg	2	2	
	XcmI	CCANNNNNNnnntgg	2	2	
50	BcgI	cgannnnntgc	3	3	
	BclI	Tgatca	3	3	
	BglI	GCCNNNNnggc	3	3	
	BsaBI	GATNNnnatc	3	3	
	BsrGI	Tgtaca	3	3	
55	SnaBI	TACgta	3	3	
	Sse8387I	CCTGCAGg	3	3	

	ApaLI	Gtgcac	4	4	LC Signal/FR1
	BspHI	Tcatga	4	4	
	BssSI	Ctcgtg	4	4	
5	PsiI	TTATAaa	4	5	
	SphI	GCATGc	4	4	
	AhdI	GACNNNnngtc	5	5	
	<b>BspEI</b>	<b>Tccgga</b>	<b>5</b>	<b>5</b>	<b>HC FR1</b>
	MscI	TGGcca	5	5	
10	SacI	GAGCTc	5	5	
	ScaI	AGTact	5	5	
	SexAI	Accwgg	5	6	
	SspI	AATatt	5	5	
	TliI	Ctcgag	5	5	
15	XhoI	Ctcgag	5	5	
	BbsI	GAAGAC	7	8	
	BstAPI	GCANNNNntgc	7	8	
	BstZ17I	GTAtac	7	7	
	EcoRV	GATatc	7	7	
	EcoRI	Gaattc	8	8	
20	BlpI	GCtnagc	9	9	
	Bsu36I	CCtnagg	9	9	
	DraIII	CACNNNgtg	9	9	
	EspI	GCtnagc	9	9	
25	StuI	AGGcct	9	13	
	<b>XbaI</b>	<b>Tctaga</b>	<b>9</b>	<b>9</b>	<b>HC FR3</b>
	<b>Bsp120I</b>	<b>Gggccc</b>	<b>10</b>	<b>11</b>	<b>CH1</b>
	<b>ApaI</b>	<b>GGGCCc</b>	<b>10</b>	<b>11</b>	<b>CH1</b>
	PspOoMI	Gggccc	10	11	
30	BciVI	GTATCCNNNNNN	11	11	
	SalI	Gtcgac	11	12	
	DrdI	GACNNNNnngtc	12	12	
	KasI	Ggcgcc	12	12	
	XmaI	Cccggg	12	14	
35	BglII	Agatct	14	14	
	HincII	GTyrac	16	18	
	BamHI	Ggatcc	17	17	
	PflMI	CCANNNNntgg	17	18	
	BsmBI	Nnnnnngagacg	18	21	
40	<b>BstXI</b>	<b>CCANNNNntgg</b>	<b>18</b>	<b>19</b>	<b>HC FR2</b>
	XmnI	GAANNnnttc	18	18	
	SacII	CCGCgg	19	19	
	PstI	CTGCAG	20	24	
	PvuII	CAGctg	20	22	
45	AvaI	Cycgrg	21	24	
	EagI	Cggccg	21	22	
	AatII	GACGTc	22	22	
	BspMI	ACCTGC	27	33	
	AccI	GTmkac	30	43	
50	StyI	Ccwwgg	36	49	
	AlwNI	CAGNNNctg	38	44	
	BsaI	GGTCTCNnnnn	38	44	
	PpuMI	RGgwccy	43	46	
	BsgI	GTGCAG	44	54	
55	BseRI	NNnnnnnnnnctcctc	48	60	
	EciI	nnnnnnnnntccgcc	52	57	
	<b>BstEII</b>	<b>Ggtnacc</b>	<b>54</b>	<b>61</b>	<b>HC Fr4, 47/79 have one</b>
	EcoO109I	RGgnccy	54	86	

BpmI    ctccag  
AvaII   Ggwcc

```

60 121
71 140

```

[illegible]

Table 5 (amended): Use of *FokI* as "Universal Restriction Enzyme"

*FokI* - for dsDNA, | represents sites of cleavage

sites of cleavage

5'-cacGGATGtg--nnnnnnn|nnnnnnn-3' (SEQ ID NO:15)

3'-gtgCCTACac--nnnnnnnnnnn|nnn-5' (SEQ ID NO:16)

RECOG

NITion of *FokI*

#### Case I

5'-...gtg|tatt-actgtgc..Substrate....-3' (SEQ ID NO:17)

3'-cac-ataa|tgacacg

gtGTAGGcac\

5'- caCATCCgtg/ (SEQ ID NO:18)

#### Case II

5'-...gtgtatt|agac-tgc..Substrate....-3' (SEQ ID NO:19)

cacataa-tctg|acg-5'

/gtgCCTACac

\cacGGATGtg-3' (SEQ ID NO:20)

#### Case III (Case I rotated 180 degrees)

/gtgCCTACac-5'

\cacGGATGtg

gtgtctt|acag-tcc-3' Adapter (SEQ ID NO:21)

3'-...cacagaa-tgtc|agg..substrate....-5' (SEQ ID NO:22)

#### Case IV (Case II rotated 180 degrees)

3'- gtGTAGGcac\ (SEQ ID NO:23)  
 5'-gag|tctc-actgagc  
 Substrate 3'-...ctc-agag|tgactcg...-5' (SEQ ID NO:24)

# Improved FokI adapters

FokI - for dsDNA, | represents sites of cleavage

## Case I

Stem 11, loop 5, stem 11, recognition 17

5'-...catgtg|tatt-actgtgc..Substrate....-3'  
 3'-gtacac-ataa|tgacacg-  
 gtGTAGGcacG T  
 5'- caCATCCgtgc C  
 TT

## Case II

Stem 10, loop 5, stem 10, recognition 18

5'-...gtgtatt|agac-tgctgcc..Substrate....-3'  
 T T  
 T gtgCCTACac  
 C cacGGATGtg-3'  
 TT

## Case III (Case I rotated 180 degrees)

Stem 11, loop 5, stem 11, recognition 20

T T  
 T TgtgCCTACac-5'  
 G AcacGGATGtg-  
 TT  
 gtgtctt|acag-tccattctg-3' Adapter  
 3'-...cacagaa-tgtc|aggtaagac..substrate....-5'

## Case IV (Case II rotated 180 degrees)

Stem 11, loop 4, stem 11, recognition 17

3'- gtGTAGGcacc T  
 caCATCCgtgg T  
 5'-atcgag|tctc-actgagc  
 Substrate 3'-...tagctc-agag|tgactcg...-5'

**BseRI**

```

                                     | sites of cleavage
5'-cacGAGGAGnnnnnnnnnn|nnnnn-3'
3'-gtgctctcnnnnnnnnn|nnnnnn-5'
      RECOG
      NITion of BseRI

```

Stem 11, loop 5, stem 11, recognition 19

```

      3'-.....gaacat|cg-ttaagccagta.....5'
┌T-T┐          cttgta-gc|aattcgggtcat-3'
C      GCTGAGGAGTC-┐
T      cgactcctcag-5'   An adapter for BseRI to cleave the substrate above.
└T┐

```

1. The first step is to identify the problem. This involves understanding the current situation and what needs to be changed.



Table 8: Matches to URE FR3 adapters in 79 human HC.

A. List of Heavy-chains genes sampled

	AF008566	af103343	HSA235676	HSU92452	HSZ93860
	AF035043	AF103367	HSA235675	HSU94412	HSZ93863
5	AF103026	AF103368	HSA235674	HSU94415	MCOMFRAA
	af103033	AF103369	HSA235673	HSU94416	MCOMFERVA
	AF103061	AF103370	HSA240559	HSU94417	S82745
	af103072	af103371	HSCB201	HSU94418	S82764
	af103078	AF103372	HSIGGVHC	HSU96389	S83240
10	AF103099	AF158381	HSU44791	HSU96391	SABVH369
	AF103102	E05213	HSU44793	HSU96392	SADEIGVH
	AF103103	E05886	HSU82771	HSU96395	SAH2IGVH
	AF103174	E05887	HSU82949	HSZ93849	SDA3IGVH
	AF103186	HSA235661	HSU82950	HSZ93850	SIGVHTTD
15	af103187	HSA235664	HSU82952	HSZ93851	SUK4IGVH
	AF103195	HSA235660	HSU82961	HSZ93853	
	af103277	HSA235659	HSU86522	HSZ93855	
	af103286	HSA235678	HSU86523	HSZ93857	
	AF103309	HSA235677			

Table 8 B. Testing all distinct GLGs from bases 89.1 to 93.2 of the heavy variable domain

	Id	Nb	0	1	2	3	4		SEQ ID NO:
	1	38	15	11	10	0	2	Seq1 gtgtattactgtgc	25
	2	19	7	6	4	2	0	Seq2 gtAtattactgtgc	26
	3	1	0	0	1	0	0	Seq3 gtgtattactgtAA	27
25	4	7	1	5	1	0	0	Seq4 gtgtattactgtAc	28
	5	0	0	0	0	0	0	Seq5 Ttgtattactgtgc	29
	6	0	0	0	0	0	0	Seq6 TtgtatCactgtgc	30
	7	3	1	0	1	1	0	Seq7 ACAattactgtgc	31
	8	2	0	2	0	0	0	Seq8 ACgtattactgtgc	32
30	9	9	2	2	4	1	0	Seq9 ATgtattactgtgc	33
	Group		26	26	21	4	2		
	Cumulative		26	52	73	77	79		

Table 8C Most important URE recognition seqs in FR3 Heavy

1	VHSzy1	GTGtattactgtgc	(ON_SHC103)	(SEQ ID NO:25)
2	VHSzy2	GTAtattactgtgc	(ON_SHC323)	(SEQ ID NO:26)
3	VHSzy4	GTGtattactgtac	(ON_SHC349)	(SEQ ID NO:28)
5	4	VHSzy9	ATGtattactgtgc	(ON_SHC5a) (SEQ ID NO:33)

Table 8D, testing 79 human HC V genes with four probes

Number of sequences..... 79  
 Number of bases..... 29143

10

		Number of mismatches								
Id	Best	0	1	2	3	4	5			
1	39	15	11	10	1	2	0	Seq1	gtgtattactgtgc	(SEQ ID NO:25)
2	22	7	6	5	3	0	1	Seq2	gtAtattactgtgc	(SEQ ID NO:26)
3	7	1	5	1	0	0	0	Seq4	gtgtattactgtAc	(SEQ ID NO:28)
4	11	2	4	4	1	0	0	Seq9	ATgtattactgtgc	(SEQ ID NO:33)
Group		25	26	20	5	2				
Cumulative		25	51	71	76	78				

15

20

One sequence has five mismatches with sequences 2, 4, and 9; it is scored as best for 2.

Id is the number of the adapter.

Best is the number of sequence for which the identified adapter was the best available.

The rest of the table shows how well the sequences match the adapters. For example, there are 11 sequences that match VHSzy1(Id=1) with 2 mismatches and are worse for all other adapters. In this sample, 90% come within 2 bases of one of the four adapters.

25

30

Table 130: PCR primers for amplification of human Ab genes

(HuIgMFOR) 5'-tgg aag agg cac gtt ctt ttc ttt-3'  
 ! (HuIgMFOREtop) 5'-aaa gaa aag aac gtg cct ctt cca-3' = reverse complement  
 (HuCKFOR) 5'-aca ctc tcc cct gtt gaa gct ctt-3'  
 (HuCL2FOR) 5'-tga aca ttc tgt agg ggc cac tg-3'  
 (HuCL7FOR) 5'-aga gca ttc tgc agg ggc cac tg-3'  
 ! Kappa  
 35 (CKForeAsc) 5'-acc gcc tcc acc ggg cgc gcc tta tta aca ctc tcc cct gtt-  
 gaa gct ctt-3'  
 (CL2ForeAsc) 5'-acc gcc tcc acc ggg cgc gcc tta tta tga aca ttc tgt-  
 agg ggc cac tg-3'  
 (CL7ForeAsc) 5'-acc gcc tcc acc ggg cgc gcc tta tta aga gca ttc tgc-  
 40 agg ggc cac tg-3'

Table 195: Human GLG FR3 sequences

45 ! VH1  
 ! 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80

agg gtc acc atg acc agg gac acg tcc atc agc aca gcc tac atg  
! 81 82 82a 82b 82c 83 84 85 86 87 88 89 90 91 92  
gag ctg agc agg ctg aga tct gac gac acg gcc gtg tat tac tgt  
! 93 94 95

5 gcg aga ga ! 1-02# 1  
aga gtc acc att acc agg gac aca tcc gcg agc aca gcc tac atg  
gag ctg agc agc ctg aga tct gaa gac acg gct gtg tat tac tgt  
gcg aga ga ! 1-03# 2

10 aga gtc acc atg acc agg aac acc tcc ata agc aca gcc tac atg  
gag ctg agc agc ctg aga tct gag gac acg gcc gtg tat tac tgt  
gcg aga gg ! 1-08# 3

aga gtc acc atg acc aca gac aca tcc acg agc aca gcc tac atg  
gag ctg agg agc ctg aga tct gac gac acg gcc gtg tat tac tgt  
gcg aga ga ! 1-18# 4

15 aga gtc acc atg acc gag gac aca tct aca gac aca gcc tac atg  
gag ctg agc agc ctg aga tct gag gac acg gcc gtg tat tac tgt  
gca aca ga ! 1-24# 5

aga gtc acc att acc agg gac agg tct atg agc aca gcc tac atg  
gag ctg agc agc ctg aga tct gag gac aca gcc atg tat tac tgt  
gca aga ta ! 1-45# 6

20 aga gtc acc atg acc agg gac acg tcc acg agc aca gtc tac atg  
gag ctg agc agc ctg aga tct gag gac acg gcc gtg tat tac tgt  
gcg aga ga ! 1-46# 7

aga gtc acc att acc agg gac atg tcc aca agc aca gcc tac atg  
gag ctg agc agc ctg aga tcc gag gac acg gcc gtg tat tac tgt  
gcg gca ga ! 1-58# 8

5      aga gtc acg att acc gcg gac gaa tcc acg agc aca gcc tac atg  
gag ctg agc agc ctg aga tct gag gac acg gcc gtg tat tac tgt  
gcg aga ga ! 1-69# 9

aga gtc acg att acc gcg gac aaa tcc acg agc aca gcc tac atg  
gag ctg agc agc ctg aga tct gag gac acg gcc gtg tat tac tgt  
gcg aga ga ! 1-e# 10

10     aga gtc acc ata acc gcg gac acg tct aca gac aca gcc tac atg  
gag ctg agc agc ctg aga tct gag gac acg gcc gtg tat tac tgt  
gca aca ga ! 1-f# 11

! VH2

15     agg ctc acc atc acc aag gac acc tcc aaa aac cag gtg gtc ctt  
aca atg acc aac atg gac cct gtg gac aca gcc aca tat tac tgt  
gca cac aga c! 2-05# 12

agg ctc acc atc tcc aag gac acc tcc aaa agc cag gtg gtc ctt  
acc atg acc aac atg gac cct gtg gac aca gcc aca tat tac tgt  
gca cgg ata c! 2-26# 13

20     agg ctc acc atc tcc aag gac acc tcc aaa aac cag gtg gtc ctt  
aca atg acc aac atg gac cct gtg gac aca gcc acg tat tac tgt  
gca cgg ata c! 2-70# 14

! VH3

25     cga ttc acc atc tcc aga gac aac gcc aag aac tca ctg tat ctg  
caa atg aac agc ctg aga gcc gag gac acg gct gtg tat tac tgt  
gcg aga ga ! 3-07# 15

cga ttc acc atc tcc aga gac aac gcc aag aac tcc ctg tat ctg  
caa atg aac agt ctg aga gct gag gac acg gcc ttg tat tac tgt  
gca aaa gat a! 3-09#16

30     cga ttc acc atc tcc agg gac aac gcc aag aac tca ctg tat ctg  
caa atg aac agc ctg aga gcc gag gac acg gcc gtg tat tac tgt  
gcg aga ga ! 3-11# 17

35     cga ttc acc atc tcc aga gaa aat gcc aag aac tcc ttg tat ctt  
caa atg aac agc ctg aga gcc ggg gac acg gct gtg tat tac tgt  
gca aga ga ! 3-13# 18

aga ttc acc atc tca aga gat gat tca aaa aac acg ctg tat ctg  
caa atg aac agc ctg aaa acc gag gac aca gcc gtg tat tac tgt  
acc aca ga ! 3-15# 19

cga ttc acc atc tcc aga gac aac gcc aag aac tcc ctg tat ctg

caa atg aac agt ctg aga gcc gag gac acg gcc ttg tat cac tgt  
gcg aga ga ! 3-20# 20

cga ttc acc atc tcc aga gac aac gcc aag aac tca ctg tat ctg  
caa atg aac agc ctg aga gcc gag gac acg gct gtg tat tac tgt  
gcg aga ga ! 3-21# 21

cgg ttc acc atc tcc aga gac aat tcc aag aac acg ctg tat ctg  
caa atg aac agc ctg aga gcc gag gac acg gcc gta tat tac tgt  
gcg aaa ga ! 3-23# 22

cga ttc acc atc tcc aga gac aat tcc aag aac acg ctg tat ctg  
caa atg aac agc ctg aga gct gag gac acg gct gtg tat tac tgt  
gcg aaa ga ! 3-30# 23

cga ttc acc atc tcc aga gac aat tcc aag aac acg ctg tat ctg  
caa atg aac agc ctg aga gct gag gac acg gct gtg tat tac tgt  
gcg aga ga ! 3303# 24

cga ttc acc atc tcc aga gac aat tcc aag aac acg ctg tat ctg  
caa atg aac agc ctg aga gct gag gac acg gct gtg tat tac tgt  
gcg aaa ga ! 3305# 25

cga ttc acc atc tcc aga gac aat tcc aag aac acg ctg tat ctg  
caa atg aac agc ctg aga gcc gag gac acg gct gtg tat tac tgt  
gcg aga ga ! 3-33# 26

cga ttc acc atc tcc aga gac aac agc aaa aac tcc ctg tat ctg  
caa atg aac agt ctg aga act gag gac acc gcc ttg tat tac tgt  
gca aaa gat a! 3-43#27

cga ttc acc atc tcc aga gac aat gcc aag aac tca ctg tat ctg  
caa atg aac agc ctg aga gac gag gac acg gct gtg tat tac tgt  
gcg aga ga ! 3-48# 28

aga ttc acc atc tca aga gat ggt tcc aaa agc atc gcc tat ctg  
caa atg aac agc ctg aaa acc gag gac aca gcc gtg tat tac tgt  
act aga ga ! 3-49# 29

cga ttc acc atc tcc aga gac aat tcc aag aac acg ctg tat ctt  
caa atg aac agc ctg aga gcc gag gac acg gcc gtg tat tac tgt  
gcg aga ga ! 3-53# 30

aga ttc acc atc tcc aga gac aat tcc aag aac acg ctg tat ctt  
caa atg ggc agc ctg aga gct gag gac atg gct gtg tat tac tgt  
gcg aga ga ! 3-64# 31

aga ttc acc atc tcc aga gac aat tcc aag aac acg ctg tat ctt  
caa atg aac agc ctg aga gct gag gac acg gct gtg tat tac tgt  
gcg aga ga ! 3-66# 32

aga ttc acc atc tca aga gat gat tca aag aac tca ctg tat ctg

caa atg aac agc ctg aaa acc gag gac acg gcc gtg tat tac tgt  
gct aga ga ! 3-72# 33

agg ttc acc atc tcc aga gat gat tca aag aac acg gcg tat ctg  
caa atg aac agc ctg aaa acc gag gac acg gcc gtg tat tac tgt  
act aga ca ! 3-73# 34

cga ttc acc atc tcc aga gac aac gcc aag aac acg ctg tat ctg  
caa atg aac agt ctg aga gcc gag gac acg gct gtg tat tac tgt  
gca aga ga ! 3-74# 35

aga ttc acc atc tcc aga gac aat tcc aag aac acg ctg cat ctt  
caa atg aac agc ctg aga gct gag gac acg gct gtg tat tac tgt  
aag aaa ga ! 3-d# 36

! VH4

cga gtc acc ata tca gta gac aag tcc aag aac cag ttc tcc ctg  
aag ctg agc tct gtg acc gcc gcg gac acg gcc gtg tat tac tgt  
gcg aga ga ! 4-04# 37

cga gtc acc atg tca gta gac acg tcc aag aac cag ttc tcc ctg  
aag ctg agc tct gtg acc gcc gtg gac acg gcc gtg tat tac tgt  
gcg aga aa ! 4-28# 38

cga gtt acc ata tca gta gac acg tct aag aac cag ttc tcc ctg  
aag ctg agc tct gtg act gcc gcg gac acg gcc gtg tat tac tgt  
gcg aga ga ! 4301# 39

cga gtc acc ata tca gta gac agg tcc aag aac cag ttc tcc ctg  
aag ctg agc tct gtg acc gcc gcg gac acg gcc gtg tat tac tgt  
gcc aga ga ! 4302# 40

cga gtt acc ata tca gta gac acg tcc aag aac cag ttc tcc ctg  
aag ctg agc tct gtg act gcc gca gac acg gcc gtg tat tac tgt  
gcc aga ga ! 4304# 41

cga gtt acc ata tca gta gac acg tct aag aac cag ttc tcc ctg  
aag ctg agc tct gtg act gcc gcg gac acg gcc gtg tat tac tgt  
gcg aga ga ! 4-31# 42

cga gtc acc ata tca gta gac acg tcc aag aac cag ttc tcc ctg  
aag ctg agc tct gtg acc gcc gcg gac acg gct gtg tat tac tgt  
gcg aga ga ! 4-34# 43

cga gtc acc ata tcc gta gac acg tcc aag aac cag ttc tcc ctg  
aag ctg agc tct gtg acc gcc gca gac acg gct gtg tat tac tgt  
gcg aga ca ! 4-39# 44

cga gtc acc ata tca gta gac acg tcc aag aac cag ttc tcc ctg  
aag ctg agc tct gtg acc gct gcg gac acg gcc gtg tat tac tgt  
gcg aga ga ! 4-59# 45

cga gtc acc ata tca gta gac acg tcc aag aac cag ttc tcc ctg  
aag ctg agc tct gtg acc gct gcg gac acg gcc gtg tat tac tgt  
gcg aga ga ! 4-61# 46

5

cga gtc acc ata tca gta gac acg tcc aag aac cag ttc tcc ctg  
aag ctg agc tct gtg acc gcc gca gac acg gcc gtg tat tac tgt  
gcg aga ga ! 4-b# 47

! VH5

10

cag gtc acc atc tca gcc gac aag tcc atc agc acc gcc tac ctg  
cag tgg agc agc ctg aag gcc tcg gac acc gcc atg tat tac tgt  
gcg aga ca ! 5-51# 48

cac gtc acc atc tca gct gac aag tcc atc agc act gcc tac ctg  
cag tgg agc agc ctg aag gcc tcg gac acc gcc atg tat tac tgt  
gcg aga ! 5-a# 49

! VH6

15

cga ata acc atc aac cca gac aca tcc aag aac cag ttc tcc ctg  
cag ctg aac tct gtg act ccc gag gac acg gct gtg tat tac tgt  
gca aga ga ! 6-1# 50

! VH7

20

cgg ttt gtc ttc tcc ttg gac acc tct gtc agc acg gca tat ctg  
cag atc tgc agc cta aag gct gag gac act gcc gtg tat tac tgt  
gcg aga ga ! 74.1# 51

---



25

Table 250: REdaptors, Extenders, and Bridges used for Cleavage and Capture of Human Heavy Chains in FR3.

A: HpyCH4V Probes of actual human HC genes

!HpyCH4V in FR3 of human HC, bases 35-56; only those with TGca site

30

TGca;10,

RE recognition:tgca

1

of length 4 is expected at 10

6-1 agttctccctgcagctgaactc

2	3-11, 3-07, 3-21, 3-72, 3-48	cactgtatctgcaaatgaacag
3	3-09, 3-43, 3-20	ccctgtatctgcaaatgaacag
4	5-51	ccgcctacctgcagtggagcag
5	3-15, 3-30, 3-30.5, 3-30.3, 3-74, 3-23, 3-33	cgctgtatctgcaaatgaacag
6	7-4.1	cggcatatctgcagatctgcag
7	3-73	cggcgtatctgcaaatgaacag
8	5-a	ctgcctacctgcagtggagcag
9	3-49	tcgcctatctgcaaatgaacag

## 10 B: HpyCH4V REaptors, Extenders, and Bridges

### B.1 REaptors

! Cutting HC lower strand:

! TmKeller for 100 mM NaCl, zero formamide

! Eaptors for cleavage

		$T_m^w$	$T_m^k$
15	(ON_HCFR36-1) 5'-agttctccccTGCAgctgaactc-3'	68.0	64.5
	(ON_HCFR36-1A) 5'-ttctccccTGCAgctgaactc-3'	62.0	62.5
	(ON_HCFR36-1B) 5'-ttctccccTGCAgctgaac-3'	56.0	59.9
	(ON_HCFR33-15) 5'-cgctgtatcTGCAaatgaacag-3'	64.0	60.8
	(ON_HCFR33-15A) 5'-ctgtatcTGCAaatgaacag-3'	56.0	56.3
20	(ON_HCFR33-15B) 5'-ctgtatcTGCAaatgaac-3'	50.0	53.1
	(ON_HCFR33-11) 5'-cactgtatcTGCAaatgaacag-3'	62.0	58.9
	(ON_HCFR35-51) 5'-ccgcctaccTGCAgtggagcag-3'	74.0	70.1

### B.2 Segment of synthetic 3-23 gene into which captured CDR3 is to be cloned

! XbaI...  
 ! D323\* cgCttcacTaag tcT aga gac aaC tcT aag aaT acT ctC taC  
 ! scab..... designed gene 3-23 gene.....

! HpyCH4V  
 ! .. .. AflIII...  
 ! Ttg caG atg aac agc TtA agG . . .  
 ! .....  
 !

### B.3 Extender and Bridges

35 ! Extender (bottom strand):

(ON\_HCHpyEx01) 5'-cAagTAGAgAgTATTcTTAgAgTTgTcTcTAGAcTTAgTgAAgcg-3'

! ON\_HCHpyEx01 is the reverse complement of

! 5'-cgCttcacTaag tcT aga gac aaC tcT aag aaT acT ctC taC Ttg -3'

40 !

! Bridges (top strand, 9-base overlap):

!  
 (ON\_HCHpyBr016-1) 5'-cgCttcacTaag tcT aqa gac aaC tcT aag-  
 aaT acT ctC taC Ttg CAgctgaac-3' {3'-term C is blocked}

5 ! 3-15 et al. + 3-11

(ON\_HCHpyBr023-15) 5'-cgCttcacTaag tcT aqa gac aaC tcT aag-  
 aaT acT ctC taC Ttg CAaatgaac-3' {3'-term C is blocked}

! 5-51

10 (ON\_HCHpyBr045-51) 5'-cgCttcacTaag tcT aqa gac aaC tcT aag-  
 aaT acT ctC taC Ttg CAgtgagc-3' {3'-term C is blocked}

! PCR primer (top strand)

15 (ON\_HCHpyPCR) 5'-cgCttcacTaag tcT aqa gac-3'

#### C: B1pI Probes from human HC GLGs

1	1-58,1-03,1-08,1-69,1-24,1-45,1-46,1-f,1-e	acatggaGCTGAGCagcctgag
2	1-02	acatggaGCTGAGCaggctgag
3	1-18	acatggagctgaggagcctgag
4	5-51,5-a	acctgcagtggagcagcctgaa
5	3-15,3-73,3-49,3-72	atctgcaaatgaacagcctgaa
6	3303,3-33,3-07,3-11,3-30,3-21,3-23,3305,3-48	atctgcaaatgaacagcctgag
7	3-20,3-74,3-09,3-43	atctgcaaatgaacagtctgag
8	74.1	atctgcagatctgcagcctaaa
9	3-66,3-13,3-53,3-d	atcttcaaatgaacagcctgag
10	3-64	atcttcaaatgggagcctgag
11	4301,4-28,4302,4-04,4304,4-31,4-34,4-39,4-59,4-61,4-b	ccctgaaGCTGAGCtctgtgac
12	6-1	ccctgcagctgaactctgtgac
13	2-70,2-05	tccttacaatgaccaacatgga
14	2-26	tccttaccatgaccaacatgga

#### D: B1pI REaptors, Extenders, and Bridges

##### 35 D.1 REaptors

		$T_m^W$	$T_m^K$
(BlpF3HC1-58)	5'-ac atg gaG CTG AGC agc ctg ag-3'	70	66.4
(BlpF3HC6-1)	5'-cc ctg aag ctg agc tct gtg ac-3'	70	66.4

! BlpF3HC6-1 matches 4-30.1, not 6-1.

40

##### D.2 Segment of synthetic 3-23 gene into which captured CDR3 is to be cloned



**F: HpyCH4III REaptors, Extenders, and Bridges****F.1 REaptors**

5	! ONs for cleavage of HC(lower) in FR3(bases 77-97)								
	! For cleavage with HpyCH4III, Bst4CI, or TaaI								
	! cleavage is in lower chain before base 88.								
	!	77	788	888	888	889	999	999	9
	!	78	901	234	567	890	123	456	7
10	(H43.77.97.1-02#1)	5'	-cc	gtg	tat	tAC	TGT	gcg	aga g-3'
								$T_m^w$	$T_m^k$
								64	62.6
	(H43.77.97.1-03#2)	5'	-ct	gtg	tat	tAC	TGT	gcg	aga g-3'
								62	60.6
	(H43.77.97.108#3)	5'	-cc	gtg	tat	tAC	TGT	gcg	aga g-3'
								64	62.6
	(H43.77.97.323#22)	5'	-cc	gtg	tat	tac	tgt	gcg	aaa g-3'
								60	58.7
	(H43.77.97.330#23)	5'	-ct	gtg	tat	tac	tgt	gcg	aaa g-3'
								60	58.7
15	(H43.77.97.439#44)	5'	-ct	gtg	tat	tac	tgt	gcg	aga g-3'
								62	60.6
	(H43.77.97.551#48)	5'	-cc	atg	tat	tac	tgt	gcg	aga g-3'
								62	60.6
	(H43.77.97.5a#49)	5'	-cc	atg	tat	tAC	TGT	gcg	aga g-3'
								58	58.3

**F.2 Extender and Bridges**

20	! XbaI and AflIII sites in bridges are bunged	
	(H43.XABr1) 5'-ggtgtagtga-	
	TCT AGt gac aac tot aag aat act ctc tac ttg cag atg -	
	aac agC TTt AGg gct qag qac aCT GCA Gtc tac tat tgt gcg aga-3'	
	(H43.XABr2) 5'-ggtgtagtga-	
25	TCT AGt gac aac tot aag aat act ctc tac ttg cag atg -	
	aac agC TTt AGg gct qag qac aCT GCA Gtc tac tat tgt gcg aaa-3'	
	(H43.XAExt) 5'-ATAgTAGAcT gcAgTgTccT cAgcccTTAA gcTgTTcATc TgcAAGTAGA-	
	gAgTATTcTT AgAgTTgTcT cTAGATcAcT AcAcc-3'	
	!H43.XAExt is the reverse complement of	
30	! 5'-ggtgtagtga-	
	!  TCT AGA gac aac tot aag aat act ctc tac ttg cag atg -	
	!  aac agC TTA AGg gct qag qac aCT GCA Gtc tac tat -3'	
	(H43.XAPCR) 5'-ggtgtagtga  TCT AGA gac aac-3'	
35	! XbaI and AflIII sites in bridges are bunged	
	(H43.ABr1) 5'-ggtgtagtga-	
	aac agC TTt AGg gct qag qac aCT GCA Gtc tac tat tgt gcg aga-3'	
	(H43.ABr2) 5'-ggtgtagtga-	
	aac agC TTt AGg gct qag qac aCT GCA Gtc tac tat tgt gcg aaa-3'	
40	(H43.AEExt) 5'-ATAgTAGAcTgcAgTgTccTcAgcccTTAAgcTgTTcAcTAcAcc-3'	

!(H43.AExt) is the reverse complement of 5'-ggtgtagtga-

! |aac|agC|TTA|AGg|gct|qag|qac|aCT|GCA|Gtc|tac|tat -3'

(H43.APCR) 5'-ggtgtagtga |aac|agC|TTA|AGg|gct|q-3'

Table 510

(FOK)act) 5'-cAATccgTg TTgTT cAcgATgTg-3'

(VHEX881) 5'-AATAGTAGAc TgcAgTgTcc TcAgccCTTA AgcTgTTcAT cTgcAAGTAG-  
AgAgTATtCT TAGAgTTgTc TcTAGAcTTA gTgAAGcg-3'

! note that VHEX881 is the reverse complement of the ON below

[RC] 5'-cgCttcacTaag-  
Scab.....  
Synthetic 3-23 as in Table 206  
|TCT|AGA|gac|aac|tct|aag|aat|act|ctc|tac|ttg|cag|atg| -  
XbaI...  
|aac|agc|TTA|AGg|gct|gag|gac|aCT|GCA|Gtc|tac|tat|t-3'  
AflII...  
5'-cgCttcacTaag-  
|TCT|AGA|gac|aac|tct|aag|aat|act|ctc|tac|ttg|cag|atg| -  
|aac|agc|TTA|AGg|gct|gag|gac|aCT|GCA|Gtc|tac|tat|tgt gcg ag-3'

(VHBA881) 5'-cgCttcacTaag-

(VHBB881) 5'-cgCttcacTaag-

(VH881PCR) 5'-cgCttcacTaaG|TCT|AGA|gac|aac-3'  
 |TCT|AGA|gac|aac|tct|aag|aat|act|ctc|tac|ttg|cag|atg|-  
 |aac|agC|TTA|AGg|ggt|gag|gac|aCT|GCA|Gtc|tac|tat|tgt|Acg|ag-3'



Table 600: V3-23 VH framework with variegated codons shown

17 18 19 20 21 22  
 A Q P A M A  
 5'-ctg tct gaa cg gcc caq cgg gcc atg gcc  
 3'-gac aga ctt gc cgg gtc ggc cgg tac cgg  
 Scab.....Sfil.....  
 Ngomi...  
 NcoI....

-----FRI(DP47/V3-23)-----  
 23 24 25 26 27 28 29 30  
 E V Q L L E S G  
 gaa|gtt|caa|ITG|tta|gag|tct|ggt|  
 gaa|gtt|caa|gtt|aac|aat|ctc|aga|cca|  
 ctt|caa|ggt|aac|aat|ctc|aga|cca|  
 | MfeI |  
 -----FRI-----  
 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45  
 G G L V Q P G G S L R L S C A  
 |ggc|ggt|ctt|gtt|caq|cgt|ggt|ggt|tct|tta|cgt|ctt|tgc|gct|  
 |ccg|cca|gaa|caa|gtc|gga|cca|cca|aga|aat|gca|gaa|aga|acg|cga|

98

53

55

50

45

40

35

Table 600: V3-23 VH framework with variegated codons shown

```

5  Sites to be varied--->      ***      ***      ***
   ---FR1----->|...CDR1.....|---FR2-----
   46 47 48 49 50 51 52 53 54 55 56 57 58 59 60
   A S G F T F S S Y A M S W V R
   |gct|TCC|GGA|ttc|act|ttc|tct|tcg|TAC|Gct|atg|tct|tgg|gtt|cgC| 143
   |cga|agg|cct|aag|tga|aag|aga|agc|atg|cga|tac|aga|acc|caa|gcg|
   | BspEI | | BsiWI | | BstXI.

10  Sites to be varies---> ***      ***      ***
   -----FR2----->|...CDR2.....
   61 62 63 64 65 66 67 68 69 70 71 72 73 74 75
   Q A P G K G L E W V S A I S G
   |CAA|gct|ccT|GGt|aaa|ggt|ttg|gag|tgg|gtt|tct|gct|atc|tct|ggt| 188
   |ggt|cga|gga|cca|ttt|cca|aac|ctc|acc|caa|aga|cga|tag|aga|cca|
15  ...BstXI |

   ***      ***
   .....CDR2.....|---FR3---
   76 77 78 79 80 81 82 83 84 85 86 87 88 89 90
   S G G S T Y Y A D S V K G R F
   |tct|ggt|ggc|agt|act|tac|tat|gct|gac|tcc|gtt|aaa|ggt|cgc|ttc| 233
   |aga|cca|ccg|tca|tga|atg|ata|cga|ctg|agg|caa|ttt|cca|gcg|aag|

25  ---FR3-----
   91 92 93 94 95 96 97 98 99 100 101 102 103 104 105
   T I S R D N S K N T L Y L Q M
   |act|atc|TCT|AGA|gac|aac|tct|aag|aat|act|ctc|tac|ttg|cag|atg| 278
   |tga|tag|aga|tct|ctg|ttg|aga|ttc|tta|tga|gag|atg|aac|gtc|tac|
   | XbaI |

30  ---FR3----->|
   106 107 108 109 110 111 112 113 114 115 116 117 118 119 120
   N S L R A E D T A V Y Y C A K
   |aac|agC|TTA|AGg|gct|gag|gac|ACT|GCA|Gtc|tac|tat|tgc|gct|aaa| 323
   |ttg|tcg|aat|tcc|cga|ctc|ctg|tga|cgt|cag|atg|ata|acg|cga|ttt|
   |AflII | | PstI |

35  .....CDR3.....|---FR4-----
   121 122 123 124 125 126 127 128 129 130 131 132 133 134 135
   D Y E G T G Y A F D I W G Q G
   |gac|tat|gaa|ggt|act|ggt|tat|gct|ttc|gaC|ATA|TGg|ggt|caa|ggt| 368
   |ctg|ata|ctt|cca|tga|cca|ata|cga|aag|ctg|tat|acc|cca|gtt|cca|
   | NdeI |

40  ---FR4----->|
   136 137 138 139 140 141 142
   T M V T V S S
   |act|atG|GTC|ACC|gtc|tct|agt- 389
   |tga|tac|cag|tgg|cag|aga|tca-
   | BstEII |

45  143 144 145 146 147 148 149 150 151 152
   A S T K G P S V F P
   gcc tcc acc aaG GGC CCa tcc GTC TTC ccc-3' 419
   cgg agg tgg ttc ccg ggt agc cag aag ggg-5'
   Bsp120I. BbsI...(2/2)
   ApaI....

50  (SFPRMET) 5'-ctg tct gaa cG GCC cag ccG-3'
   (TOPFR1A) 5'-ctg tct gaa cG GCC cag ccG GCC atg gcc-
   gaa|gtt|CAA|TTG|tta|gag|tct|ggt|-
   |ggc|ggt|ctt|gtt|cag|cct|ggt|ggt|tct|tta-3'
   (BOTFR1B) 3'-caa|gtc|gga|cca|cca|aga|aat|gca|gaa|aga|acg|cga|-
   |cga|agg|cct|aag|tga|aag-5' ! bottom strand

```

```

(BOTFR2)  3'-acc|caa|gcg|-
           |gtt|cga|gga|cca|ttt|cca|aac|ctc|acc|caa|aga|-5' ! bottom strand
(BOTFR3)  3'- a|cga|ctg|agg|caa|ttt|cca|gcg|aag|-
           |tga|tag|aga|tct|ctg|ttg|aga|ttc|tta|tga|gag|atg|aac|gtc|tac|-
5         |ttg|tcg|aat|tcc|cga|ctc|ctg|tga-5'
(F06)     5'-gC|TTA|AGg|gct|gag|gac|aCT|GCA|Gtc|tac|tat|tgc|gct|aaa|-
           |gac|tat|gaa|ggt|act|ggt|tat|gct|ttc|gaC|ATA|TGg|ggt|c-3'
(BOTFR4)  3'-cga|aag|ctg|tat|acc|cca|gtt|cca|-
           |tga|tac|cag|tgg|cag|aga|tca-
10        cgg agg tgg ttc ccg ggt agc cag aag ggg-5' ! bottom strand
(BOTPRCPRIM)  3'-gg ttc ccg ggt agc cag aag ggg-5'
! CDR1 diversity
!
15 (ON-vgC1)  5'-gct|TCC|GGA|ttc|act|tct|<1>|TAC|<1>|atg|<1>|-
           ! CDR1.....6859
           |tgg|ggt|cgc|CAa|gct|ccT|GG-3'
!
! <1> stands for an equimolar mix of {ADEF GHIKLMNPQRSTVWY}; no C
20 ! (this is not a sequence)
! CDR2 diversity
!
25 (ON-vgC2)  5'-ggt|ttg|gag|tgg|ggt|tct|<2>|atc|<2>|<3>|-
           ! CDR2.....
           |tct|ggt|ggc|<1>|act|<1>|tat|gct|gac|tcc|gtt|aaa|gg-3'
           CDR2.....
! <1> is an equimolar mixture of {ADEF GHIKLMNPQRSTVWY}; no C
! <2> is an equimolar mixture of {YRWVGS}; no ACDEFHIKLMNPQT
30 ! <3> is an equimolar mixture of {PS}; no ACDEF GHIKLMNPQRSTVWY

```

Table 800 (new)

The following list of enzymes was taken from  
<http://rebase.neb.com/cgi-bin/asymmlist>.

I have removed the enzymes that a) cut within the recognition, b) cut on both sides of the recognition, or c) have fewer than 2 bases between recognition and closest cut site.

REBASE Enzymes  
 04/13/2001

Type II restriction enzymes with asymmetric recognition sequences:

Enzymes	Recognition Sequence	Isoschizomers	Suppliers
AarI	CACCTGCNNNN^NNNN	-	Y
AceIII	CAGCTCNNNNNNN^NNNN	-	-
Bbr7I	GAAGACNNNNNNN^NNNN	-	-
BbvI	GCAGCNNNNNNNN^NNNN	-	Y
BbvII	GAAGACNN^NNNN	-	-
Bce83I	CTTGAGNNNNNNNNNNNNNN^NN	-	-
BceAI	ACGGCNNNNNNNNNNNN^NN	-	Y
BceFI	ACGGCNNNNNNNNNNNN^N	-	-
BciVI	GTATCCNNNNN N^	BfuI	Y
BfiI	ACTGGGNNNN N^	BmrI	Y
BinI	GGATCNNNN^N	-	-
BscAI	GCATCNNNN^NN	-	-
BseRI	GAGGAGNNNNNNNN NN^	-	Y
BsmFI	GGGACNNNNNNNNNN^NNNN	BspLU11III	Y
BspMI	ACCTGCNNNN^NNNN	Acc36I	Y
EciI	GGCGGANNNNNNNNNN NN^	-	Y
Eco57I	CTGAAGNNNNNNNNNNNNNN NN^	BspKT5I	Y
FauI	CCCGCNNNN^NN	BstFZ438I	Y
FokI	GGATGNNNNNNNNNN^NNNN	BstPZ418I	Y
GsuI	CTGGAGNNNNNNNNNNNNNN NN^	-	Y
HgaI	GACGCNNNNN^NNNNN	-	Y
HphI	GGTGANNNNNNN N^	AsuHPI	Y
MboII	GAAGANNNNNNN N^	-	Y
MlyI	GAGTCNNNNN^	SchI	Y
MmeI	TCCRACNNNNNNNNNNNNNNNN NN^	-	-
MnlI	CCTCNNNNNN N^	-	Y
PleI	GAGTCNNNN^N	PpsI	Y
RleAI	CCCACANNNNNNNNNN NNN^	-	-
SfaNI	GCATCNNNNN^NNNN	BspST5I	Y
SspD5I	GGTGANNNNNNNN^	-	-
Sth132I	CCCGNNNN^NNNN	-	-
StsI	GGATGNNNNNNNNNN^NNNN	-	-
TaqII	GACCGANNNNNNNNNN NN^, CACCCANNNNNNNNNN NN^	-	-
Tth111III	CAARCANNNNNNNNN NN^	-	-
UbaPI	CGAACG	-	-

The notation is  $\wedge$  means cut the upper strand and  $\_$  means cut the lower strand. If the upper and lower strand are cut at the same place, then only  $\wedge$  appears.

11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000  
1001  
1002  
1003  
1004  
1005  
1006  
1007  
1008  
1009  
1010  
1011  
1012  
1013  
1014  
1015  
1016  
1017  
1018  
1019  
1020  
1021  
1022  
1023  
1024  
1025  
1026  
1027  
1028  
1029  
1030  
1031  
1032  
1033  
1034  
1035  
1036  
1037  
1038  
1039  
1040  
1041  
1042  
1043  
1044  
1045  
1046  
1047  
1048  
1049  
1050  
1051  
1052  
1053  
1054  
1055  
1056  
1057  
1058  
1059  
1060  
1061  
1062  
1063  
1064  
1065  
1066  
1067  
1068  
1069  
1070  
1071  
1072  
1073  
1074  
1075  
1076  
1077  
1078  
1079  
1080  
1081  
1082  
1083  
1084  
1085  
1086  
1087  
1088  
1089  
1090  
1091  
1092  
1093  
1094  
1095  
1096  
1097  
1098  
1099  
1100  
1101  
1102  
1103  
1104  
1105  
1106  
1107  
1108  
1109  
1110  
1111  
1112  
1113  
1114  
1115  
1116  
1117  
1118  
1119  
1120  
1121  
1122  
1123  
1124  
1125  
1126  
1127  
1128  
1129  
1130  
1131  
1132  
1133  
1134  
1135  
1136  
1137  
1138  
1139  
1140  
1141  
1142  
1143  
1144  
1145  
1146  
1147  
1148  
1149  
1150  
1151  
1152  
1153  
1154  
1155  
1156  
1157  
1158  
1159  
1160  
1161  
1162  
1163  
1164  
1165  
1166  
1167  
1168  
1169  
1170  
1171  
1172  
1173  
1174  
1175  
1176  
1177  
1178  
1179  
1180  
1181  
1182  
1183  
1184  
1185  
1186  
1187  
1188  
1189  
1190  
1191  
1192  
1193  
1194  
1195  
1196  
1197  
1198  
1199  
1200  
1201  
1202  
1203  
1204  
1205  
1206  
1207  
1208  
1209  
1210  
1211  
1212  
1213  
1214  
1215  
1216  
1217  
1218  
1219  
1220  
1221  
1222  
1223  
1224  
1225  
1226  
1227  
1228  
1229  
1230  
1231  
1232  
1233  
1234  
1235  
1236  
1237  
1238  
1239  
1240  
1241  
1242  
1243  
1244  
1245  
1246  
1247  
1248  
1249  
1250  
1251  
1252  
1253  
1254  
1255  
1256  
1257  
1258  
1259  
1260  
1261  
1262  
1263  
1264  
1265  
1266  
1267  
1268  
1269  
1270  
1271  
1272  
1273  
1274  
1275  
1276  
1277  
1278  
1279  
1280  
1281  
1282  
1283  
1284  
1285  
1286  
1287  
1288  
1289  
1290  
1291  
1292  
1293  
1294  
1295  
1296  
1297  
1298  
1299  
1300  
1301  
1302  
1303  
1304  
1305  
1306  
1307  
1308  
1309  
1310  
1311  
1312  
1313  
1314  
1315  
1316  
1317  
1318  
1319  
1320  
1321  
1322  
1323  
1324  
1325  
1326  
1327  
1328  
1329  
1330  
1331  
1332  
1333  
1334  
1335  
1336  
1337  
1338  
1339  
1340  
1341  
1342  
1343  
1344  
1345  
1346  
1347  
1348  
1349  
1350  
1351  
1352  
1353  
1354  
1355  
1356  
1357  
1358  
1359  
1360  
1361  
1362  
1363  
1364  
1365  
1366  
1367  
1368  
1369  
1370  
1371  
1372  
1373  
1374  
1375  
1376  
1377  
1378  
1379  
1380  
1381  
1382  
1383  
1384  
1385  
1386  
1387  
1388  
1389  
1390  
1391  
1392  
1393  
1394  
1395  
1396  
1397  
1398  
1399  
1400  
1401  
1402  
1403  
1404  
1405  
1406  
1407  
1408  
1409  
1410  
1411  
1412  
1413  
1414  
1415  
1416  
1417  
1418  
1419  
1420  
1421  
1422  
1423  
1424  
1425  
1426  
1427  
1428  
1429  
1430  
1431  
1432  
1433  
1434  
1435  
1436  
1437  
1438  
1439  
1440  
1441  
1442  
1443  
1444  
1445  
1446  
1447  
1448  
1449  
1450  
1451  
1452  
1453  
1454  
1455  
1456  
1457  
1458  
1459  
1460  
1461  
1462  
1463  
1464  
1465  
1466  
1467  
1468  
1469  
1470  
1471  
1472  
1473  
1474  
1475  
1476  
1477  
1478  
1479  
1480  
1481  
1482  
1483  
1484  
1485  
1486  
1487  
1488  
1489  
1490  
1491  
1492  
1493  
1494  
1495  
1496  
1497  
1498  
1499  
1500  
1501  
1502  
1503  
1504  
1505  
1506  
1507  
1508  
1509  
1510  
1511  
1512  
1513  
1514  
1515  
1516  
1517  
1518  
1519  
1520  
1521  
1522  
1523  
1524  
1525  
1526  
1527  
1528  
1529  
1530  
1531  
1532  
1533  
1534  
1535  
1536  
1537  
1538  
1539  
1540  
1541  
1542  
1543  
1544  
1545  
1546  
1547  
1548  
1549  
1550  
1551  
1552  
1553  
1554  
1555  
1556  
1557  
1558  
1559  
1560  
1561  
1562  
1563  
1564  
1565  
1566  
1567  
1568  
1569  
1570  
1571  
1572  
1573  
1574  
1575  
1576  
1577  
1578  
1579  
1580  
1581  
1582  
1583  
1584  
1585  
1586  
1587  
1588  
1589  
1590  
1591  
1592  
1593  
1594  
1595  
1596  
1597  
1598  
1599  
1600  
1601  
1602  
1603  
1604  
1605  
1606  
1607  
1608  
1609  
1610  
1611  
1612  
1613  
1614  
1615  
1616  
1617  
1618  
1619  
1620  
1621  
1622  
1623  
1624  
1625  
1626  
1627  
1628  
1629  
1630  
1631  
1632  
1633  
1634  
1635  
1636  
1637  
1638  
1639  
1640  
1641  
1642  
1643  
1644  
1645  
1646  
1647  
1648  
1649  
1650  
1651  
1652  
1653  
1654  
1655  
1656  
1657  
1658  
1659  
1660  
1661  
1662  
1663  
1664  
1665  
1666  
1667  
1668  
1669  
1670  
1671  
1672  
1673  
1674  
1675  
1676  
1677  
1678  
1679  
1680  
1681  
1682  
1683  
1684  
1685  
1686  
1687  
1688  
1689  
1690  
1691  
1692  
1693  
1694  
1695  
1696  
1697  
1698  
1699  
1700  
1701  
1702  
1703  
1704  
1705  
1706  
1707  
1708  
1709  
1710  
1711  
1712  
1713  
1714  
1715  
1716  
1717  
1718  
1719  
1720  
1721  
1722  
1723  
1724  
1725  
1726  
1727  
1728  
1729  
1730  
1731  
1732  
1733  
1734  
1735  
1736  
1737  
1738  
1739  
1740  
1741  
1742  
1743  
1744  
1745  
1746  
1747  
1748  
1749  
1750  
1751  
1752  
1753  
1754  
1755  
1756  
1757  
1758  
1759  
1760  
1761  
1762  
1763  
1764  
1765  
1766  
1767  
1768  
1769  
1770  
1771  
1772  
1773  
1774  
1775  
1776  
1777  
1778  
1779  
1780  
1781  
1782  
1783  
1784  
1785  
1786  
1787  
1788  
1789  
1790  
1791  
1792  
1793  
1794  
1795  
1796  
1797  
1798  
1799  
1800  
1801  
1802  
1803  
1804  
1805  
1806  
1807  
1808  
1809  
1810  
1811  
1812  
1813  
1814  
1815  
1816  
1817  
1818  
1819  
1820  
1821  
1822  
1823  
1824  
1825  
1826  
1827  
1828  
1829  
1830  
1831  
1832  
1833  
1834  
1835  
1836  
1837  
1838  
1839  
1840  
1841  
1842  
1843  
1844  
1845  
1846  
1847  
1848  
1849  
1850  
1851  
1852  
1853  
1854  
1855  
1856  
1857  
1858  
1859  
1860  
1861  
1862  
1863  
1864  
1865  
1866  
1867  
1868  
1869  
1870  
1871  
1872  
1873  
1874  
1875  
1876  
1877  
1878  
1879  
1880  
1881  
1882  
1883  
1884  
1885  
1886  
1887  
1888  
1889  
1890  
1891  
1892  
1893  
1894  
1895  
1896  
1897  
1898  
1899  
1900  
1901  
1902  
1903  
1904  
1905  
1906  
1907  
1908  
1909  
1910  
1911  
1912  
1913  
1914  
1915  
1916  
1917  
1918  
1919  
1920  
1921  
1922  
1923  
1924  
1925  
1926  
1927  
1928  
1929  
1930  
1931  
1932  
1933  
1934  
1935  
1936  
1937  
1938  
1939  
1940  
1941  
1942  
1943  
1944  
1945  
1946  
1947  
1948  
1949  
1950  
1951  
1952  
1953  
1954  
1955  
1956  
1957  
1958  
1959  
1960  
1961  
1962  
1963  
1964  
1965  
1966  
1967  
1968  
1969  
1970  
1971  
1972  
1973  
1974  
1975  
1976  
1977  
1978  
1979  
1980  
1981  
1982  
1983  
1984  
1985  
1986  
1987  
1988  
1989  
1990  
1991  
1992  
1993  
1994  
1995  
1996  
1997  
1998  
1999  
2000  
2001  
2002  
2003  
2004  
2005  
2006  
2007  
2008  
2009  
2010  
2011  
2012  
2013  
2014  
2015  
2016  
2017  
2018  
2019  
2020  
2021  
2022  
2023  
2024  
2025  
2026  
2027  
2028  
2029  
2030  
2031  
2032  
2033  
2034  
2035  
2036  
2037  
2038  
2039  
2040  
2041  
2042  
2043  
2044  
2045  
2046  
2047  
2048  
2049  
2050  
2051  
2052  
2053  
2054  
2055  
2056  
2057  
2058  
2059  
2060  
2061  
2062  
2063  
2064  
2065  
2066  
2067  
2068  
2069  
2070  
2071  
2072  
2073  
2074  
2075  
2076  
2077  
2078  
2079  
2080  
2081  
2082  
2083  
2084  
2085  
2086  
2087  
2088  
2089  
2090  
2091  
2092  
2093  
2094  
2095  
2096  
2097  
2098  
2099  
2100  
2101  
2102  
2103  
2104  
2105  
2106  
2107  
2108  
2109  
2110  
2111  
2112  
2113  
2114  
2115  
2116  
2117  
2118  
2119  
2120  
2121  
2122  
2123  
2124  
2125  
2126  
2127  
2128  
2129  
2130  
2131  
2132  
2133  
2134  
2135  
2136  
2137  
2138  
2139  
2140  
2141  
2142  
2143  
2144  
2145  
2146  
2147  
2148  
2149  
2150  
2151  
2152  
2153  
2154  
2155  
2156  
2157  
2158  
2159  
2160  
2161  
2162  
2163  
2164  
2165  
2166  
2167  
2168  
2169  
2170  
2171  
2172  
2173  
2174  
2175  
2176  
2177  
2178  
2179  
2180  
2181  
2182  
2183  
2184  
2185  
2186  
2187  
2188  
2189  
2190  
2191  
2192  
2193  
2194  
2195  
2196  
2197  
2198  
2199  
2200  
2201  
2202  
2203  
2204  
2205  
2206  
2207  
2208  
2209  
2210  
2211  
2212  
2213  
2214  
2215  
2216  
2217  
2218  
2219  
2220  
2221  
2222  
2223  
2224  
2225  
2226  
2227  
2228  
2229  
2230  
2231  
2232  
2233  
2234  
2235

Table 120: MALIA3, annotated

! MALIA3 9532 bases

!-----

```

1 aat gct act act att agt aga att gat gcc acc ttt tca gct cgc gcc
5 ! gene ii continued
  49 cca aat gaa aat ata gct aaa cag gtt att gac cat ttg cga aat gta
  97 tct aat ggt caa act aaa tct act cgt tcg cag aat tgg gaa tca _act
 145 gtt aca tgg aat gaa act tcc aga cac cgt act tta gtt gca tat tta
 193 aaa cat gtt gag cta cag cac cag att cag caa tta agc tct aag cca
10 241 tcc gca aaa atg acc tct tat caa aag gag caa tta aag gta ctc tct
 289 aat cct gac ctg ttg gag ttt gct tcc ggt ctg gtt cgc ttt gaa gct
 337 cga att aaa acg cga tat ttg aag tct ttc ggg ctt cct ctt aat ctt
 385 ttt gat gca atc cgc ttt gct tct gac tat aat agt cag ggt aaa gac
 433 ctg att ttt gat tta tgg tca ttc tcg ttt tct gaa ctg ttt aaa gca
15 481 ttt gag ggg gat tca ATG aat att tat gac gat tcc gca gta ttg gac
!      RBS?.....      Start gene x, ii continues
 529 gct atc cag tct aaa cat ttt act att acc ccc tct ggc aaa act tct
 577 ttt gca aaa gcc tct cgc tat ttt ggt ttt tat cgt cgt ctg gta aac
 625 gag ggt tat gat agt gtt gct ctt act atg cct cgt aat tcc ttt tgg
20 673 cgt tat gta tct gca tta gtt gaa tgt ggt att cct aaa tct caa ctg
 721 atg aat ctt tct acc tgt aat aat gtt gtt ccg tta gtt cgt ttt att
 769 aac gta gat ttt tct tcc caa cgt cct gac tgg tat aat gag cca gtt
 817 ctt aaa atc gca TAA
!
!      End X & II
25 832 ggtaattca ca
!
!      M1      E5      Q10      T15
843 ATG att aaa gtt gaa att aaa cca tct caa gcc caa ttt act act cgt
!      Start gene V
30 !
!      S17      S20      P25      E30
891 tct ggt gtt tct cgt cag ggc aag cct tat tca ctg aat gag cag ctt
!
!      V35      E40      V45
35 939 tgt tac gtt gat ttg ggt aat gaa tat ccg gtt ctt gtc aag att act
!
!      D50      A55      L60
987 ctt gat gaa ggt cag cca gcc tat gcg cct ggt cTG TAC Acc gtt cat
!      BsrGI...

```

```

!           L65                V70                S75                R80
1035 ctg tcc tct ttc aaa gtt ggt cag ttc ggt tcc ctt atg att gac cgt
!
!
!           P85      K87 end of V
5  1083 ctg cgc ctc gtt ccg gct aag TAA C
!
!
1108 ATG gag cag gtc gcg gat ttc gac aca att tat cag gcg atg
!
!   Start gene VII
!
10 1150 ata caa atc tcc gtt gta ctt tgt ttc gcg ctt ggt ata atc
!
!
!           VII and IX overlap.
!           ..... S2  V3  L4  V5                S10
1192 gct ggg ggt caa agA TGA gt gtt tta gtg tat tct ttc gcc tct ttc gtt
15 1
!           End VII
!           |start IX
!
!           L13      W15                G20                T25                E29
1242 tta ggt tgg tgc ctt cgt agt ggc att acg tat ttt acc cgt tta atg gaa
!
!
20 1293 act tcc tc
!
!
!           .... stop of IX, IX and VIII overlap by four bases
1301 ATG aaa aag tct tta gtc ctc aaa gcc tct gta gcc gtt gct acc ctc
!
!   Start signal sequence of viii.
25 1
!
1349 gtt ccg atg ctg tct ttc gct gct gag ggt gac gat ccc gca aaa gcg
!
!           mature VIII --->
1397 gcc ttt aac tcc ctg caa gcc tca gcg acc gaa tat atc ggt tat gcg
1445 tgg gcg atg gtt gtt gtc att
30 1466 gtc ggc gca act atc ggt atc aag ctg ttt aag
1499 aaa ttc acc tcg aaa gca ! 1515
!
!           ..... -35 ..
!
1517      agc tga taaaccgat acaattaaag gctccttttg
35 1
!           ..... -10 ...
!
1552 gageccttttt ttttGGAGAt ttt ! S.D. underlined
!
!           <----- III signal sequence ----->

```

! M K K L L F A I P L V  
1575 caac GTG aaa aaa tta tta ttc gca att cct tta gtt ! 1611

! V P F Y S H S A Q  
5 1612 gtt cct ttc tat tct cac aGT gcA Cag tCT  
! ApaLI...  
!

1642 GTC GTG ACG CAG CCG CCC TCA GTG TCT GGG GCC CCA GGG CAG  
AGG GTC ACC ATC TCC TGC ACT GGG AGC AGC TCC AAC ATC GGG GCA  
10 ! BstEII...

1729 GGT TAT GAT GTA CAC TGG TAC CAG CAG CTT CCA GGA ACA GCC CCC AAA  
1777 CTC CTC ATC TAT GGT AAC AGC AAT CGG CCC TCA GGG GTC CCT GAC CGA  
1825 TTC TCT GGC TCC AAG TCT GGC ACC TCA GCC TCC CTG GCC ATC ACT  
1870 GGG CTC CAG GCT GAG GAT GAG GCT GAT TAT  
15 1900 TAC TGC CAG TCC TAT GAC AGC AGC CTG AGT  
1930 GGC CTT TAT GTC TTC GGA ACT GGG ACC AAG GTC ACC GTC

! BstEII...  
1969 CTA GGT CAG CCC AAG GCC AAC CCC ACT GTC ACT  
2002 CTG TTC CCG CCC TCC TCT GAG GAG CTC CAA GCC AAC AAG GCC ACA CTA  
20 2050 GTG TGT CTG ATC AGT GAC TTC TAC CCG GGA GCT GTG ACA GTG GCC TGG  
2098 AAG GCA GAT AGC AGC CCC GTC AAG GCG GGA GTG GAG ACC ACC ACA CCC  
2146 TCC AAA CAA AGC AAC AAC AAG TAC GCG GCC AGC AGC TAT CTG AGC CTG  
2194 ACG CCT GAG CAG TGG AAG TCC CAC AGA AGC TAC AGC TGC CAG GTC ACG  
2242 CAT GAA GGG AGC ACC GTG GAG AAG ACA GTG GCC CCT ACA GAA TGT TCA  
25 2290 TAA TAA ACCG CCTCCACCGG GCGCGCCAAT TCTATTTCAA GGAGACAGTC ATA

! AscI.....  
!

! PelB signal----->  
! M K Y L L P T A A A G L L L L  
30 2343 ATG AAA TAC CTA TTG CCT ACG GCA GCC GCT GGA TTG TTA TTA CTC  
!

! 16 17 18 19 20 21 22  
! A A Q P A M A  
2388 gcG GCC cag ccG GCC atg gcc

35 ! SfiI.....

! NgoMI...(1/2)

! NcoI.....  
!



```

5      !      FR1(DP47/V3-23)-----
      !      23  24  25  26  27  28  29  30
      !      E   V   Q   L   L   E   S   G
2409   !      gaa|ggt|CAA|TTG|tta|gag|tct|ggt|
      !      | MfeI |
      !
      !      -----FR1-----
      !      31  32  33  34  35  36  37  38  39  40  41  42  43  44  45
      !      G   G   L   V   Q   P   G   G   S   L   R   L   S   C   A
10    2433 |ggc|ggt|ctt|ggt|cag|cct|ggt|ggt|tct|tta|cgt|ctt|tct|tgc|gct|
      !
      !      ----FR1----->|...CDR1.....|---FR2-----
      !      46  47  48  49  50  51  52  53  54  55  56  57  58  59  60
      !      A   S   G   F   T   F   S   S   Y   A   M   S   W   V   R
15    2478 |gct|TCC|GGA|ttc|act|ttc|tct|tCG|TAC|Gct|atg|tct|tgg|ggt|cgC|
      !      | BspEI |                      | BsiWI|                      |BstXI.
      !
      !      -----FR2----->|...CDR2.....
      !      61  62  63  64  65  66  67  68  69  70  71  72  73  74  75
      !      Q   A   P   G   K   G   L   E   W   V   S   A   I   S   G
20    2523 |CAa|gct|cct|GGt|aaa|ggt|ttg|gag|tgg|ggt|tct|gct|atc|tct|ggt|
      !      ...BstXI                      |
      !
      !      .....CDR2.....|---FR3---
25    !      76  77  78  79  80  81  82  83  84  85  86  87  88  89  90
      !      S   G   G   S   T   Y   Y   A   D   S   V   K   G   R   F
      !      2568 |tct|ggt|ggc|agt|act|tac|tat|gct|gac|tcc|ggt|aaa|ggt|cgc|ttc|
      !
      !
30    !      -----FR3-----
      !      91  92  93  94  95  96  97  98  99 100 101 102 103 104 105
      !      T   I   S   R   D   N   S   K   N   T   L   Y   L   Q   M
      !      2613 |act|atc|TCT|AGA|gac|aac|tct|aag|aat|act|ctc|tac|ttg|cag|atg|
      !      | XbaI |
35    !
      !      ---FR3----->|
      !      106 107 108 109 110 111 112 113 114 115 116 117 118 119 120
      !      N   S   L   R   A   E   D   T   A   V   Y   Y   C   A   K
      !      2658 |aac|agC|TTA|AGg|gct|gag|gac|aCT|GCA|Gtc|tac|tat|tgc|gct|aaa|

```

```

!               |AflIII |               | PstI |
!
!       .....CDR3.....|----FR4-----
!       121 122 123 124 125 126 127 128 129 130 131 132 133 134 135
5      !       D   Y   E   G   T   G   Y   A   F   D   I   W   G   Q   G
      2703 |gac|tat|gaa|ggt|act|ggt|tat|gct|ttc|gaC|ATA|TGg|ggt|caa|ggt|
!               | NdeI | (1/4)
!
!       -----FR4----->|
10     !       136 137 138 139 140 141 142
!       T   M   V   T   V   S   S
      2748 |act|atG|GTC|ACC|gtc|tct|agt
!               | BstEII |
! From BstEII onwards, pV323 is same as pCES1, except as noted.
15     ! BstEII sites may occur in light chains; not likely to be unique in final
!       vector.
!
!               143 144 145 146 147 148 149 150 151 152
!               A   S   T   K   G   P   S   V   F   P
20     !       2769 gcc tcc acc aaG GGC CCa tcg GTC TTC ccc
!               Bsp120I.      BbsI... (2/2)
!               ApaI....
!
!       153 154 155 156 157 158 159 160 161 162 163 164 165 166 167
25     !       L   A   P   S   S   K   S   T   S   G   G   T   A   A   L
      2799 ctg gca ccC TCC TCc aag agc acc tct ggg ggc aca gcg gcc ctg
!               BseRI... (2/2)
!
!       168 169 170 171 172 173 174 175 176 177 178 179 180 181 182
30     !       G   C   L   V   K   D   Y   F   P   E   P   V   T   V   S
      2844 ggc tgc ctg GTC AAG GAC TAC TTC CCc gaA CCG GTg acg gtg tcg
!               AgeI....
!
!       183 184 185 186 187 188 189 190 191 192 193 194 195 196 197
35     !       W   N   S   G   A   L   T   S   G   V   H   T   F   P   A
      2889 tgg aac tca GGC GCC ctg acc agc ggc gtc cac acc ttc ccg gct
!               KasI... (1/4)
!
!       198 199 200 201 202 203 204 205 206 207 208 209 210 211 212

```

```

!       V   L   Q   S   S   G   L   Y   S   L   S   S   V   V   T
2934   gtc cta cag tCt agc GGa ctc tac tcc ctc agc agc gta gtg acc
!       (Bsu36I...) (knocked out)
!
5   !       213 214 215 216 217 218 219 220 221 222 223 224 225 226 227
!       V   P   S   S   S   L   G   T   Q   T   Y   I   C   N   V
2979   gtg ccC tCt tct agc tTG Ggc acc cag acc tac atc tgc aac gtg
!       (BstXI.....)N.B. destruction of BstXI & BpmI sites.
!
10  !       228 229 230 231 232 233 234 235 236 237 238 239 240 241 242
!       N   H   K   P   S   N   T   K   V   D   K   K   V   E   P
3024   aat cac aag ccc agc aac acc aag gtg gac aag aaa gtt gag ccc
!
!       243 244 245
15  !       K   S   C   A   A   A   H   H   H   H   H   H   S   A
3069   aaa tct tgt GCG GCC Gct cat cac cac cat cat cac tct gct
!       NotI.....
!
!       E   Q   K   L   I   S   E   E   D   L   N   G   A   A
20  !       3111   gaa caa aaa ctc atc tca gaa gag gat ctg aat ggt gcc gca
!
!       D   I   N   D   D   R   M   A   S   G   A
!       3153   GAT ATC aac gat gat cgt atg   gct AGC   ggc gcc
25  !       rEK cleavage site.....   NheI...   KasI...
!       EcoRV..
!
!   Domain 1 -----
!       A   E   T   V   E   S   C   L   A
30  !       3183   gct gaa act gtt gaa agt tgt tta gca
!
!       K   P   H   T   E   I   S   F
!       3210   aaa ccc cat aca gaa aat tca ttt
35  !
!       T   N   V   W   K   D   D   K   T
!       3234   aCT AAC GTC TGG AAA GAC GAC AAA ACT
!
!       L   D   R   Y   A   N   Y   E   G   C   L   W   N   A   T   G   V

```

```

3261 tta gat cgt tac gct aac tat gag ggt tgt ctg tgG AAT GcT aca ggc gtt
!
! BsmI_____
!
! V V C T G D E T Q C Y G T W V P I
5 3312 gta gtt tgt act ggt GAC GAA ACT CAG TGT TAC GGT ACA TGG GTT cct att
!
! G L A I P E N
3363 ggg ctt gct atc cct gaa aat
!
10 ! L1 linker -----
! E G G G S E G G G S
3384 gag ggt ggt ggc tct gag ggt ggc ggt tct
!
! E G G G S E G G G T
15 3414 gag ggt ggc ggt tct gag ggt ggc ggt act
!
! Domain 2 -----
3444 aaa cct cct gag tac ggt gat aca cct att ccg ggc tat act tat atc aac
3495 cct ctc gac ggc act tat ccg cct ggt act gag caa aac ccc gct aat cct
20 3546 aat cct tct ctt GAG GAG tct cag cct ctt aat act ttc atg ttt cag aat
!
! BseRI_____
3597 aat agg ttc cga aat agg cag ggg gca tta act gtt tat acg ggc act
3645 gtt act caa ggc act gac ccc gtt aaa act tat tac cag tac act cct
3693 gta tca tca aaa gcc atg tat gac gct tac tgg aac ggt aaa ttc AGA
25 !
! AlwNI
3741 GAC TGc gct ttc cat tct ggc ttt aat gaa gat cca ttc gtt tgt gaa
!
! AlwNI
3789 tat caa ggc caa tcg tct gac ctg cct caa cct cct gtc aat gct
!
30 3834 ggc ggc ggc tct
! start L2 -----
3846 ggt ggt ggt tct
3858 ggt ggc ggc tct
3870 gag ggt ggt ggc tct gag ggt ggc ggt tct
35 3900 gag ggt ggc ggc tct gag gga ggc ggt tcc
3930 ggt ggt ggc tct ggt ! end L2
!
! Domain 3 -----
! S G D F D Y E K M A N A N K G A

```



4451 ttc ctt ctg gta act ttg ttc ggc tat ctg ctt act ttt ctt aaa aag  
 4499 ggc ttc ggt aag ata gct att gct att tca ttg ttt ctt gct ctt att  
 4547 att ggg ctt aac tca att ctt gtg ggt tat ctc tct gat att agc gct  
 4595 caa tta ccc tct gac ttt gtt cag ggt gtt cag tta att ctc ccg tct  
 5 4643 aat gcg ctt ccc tgt ttt tat gtt att ctc tct gta aag gct gct att  
 4691 ttc att ttt gac gtt aaa caa aaa atc gtt tct tat ttg gat tgg gat

!

!

M1 A2 V3 F5 L10 G13

10 4739 aaa TAA t ATG gct gtt tat ttt gta act ggc aaa tta ggc tct gga  
 ! end VI Start gene I

!

! 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28  
 ! K T L V S V G K I Q D K I V A

15 4785 aag acg ctc gtt agc gtt ggt aag att cag gat aaa att gta gct

!

! 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43  
 ! G C K I A T N L D L R L Q N L

4830 ggg tgc aaa ata gca act aat ctt gat tta agg ctt caa aac ctc

!

20 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58  
 ! P Q V G R F A K T P R V L R I

4875 ccg caa gtc ggg agg ttc gct aaa acg cct cgc gtt ctt aga ata

!

25 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73  
 ! P D K P S I S D L L A I G R G

4920 ccg gat aag cct tct ata tct gat ttg ctt gct att ggg cgc ggt

!

30 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88  
 ! N D S Y D E N K N G L L V L D

4965 aat gat tcc tac gat gaa aat aaa aac ggc ttg ctt gtt ctc gat

!

89 90 91 92 93 94 95 96 97 98 99 100 101 102 103  
 ! E C G T W F N T R S W N D K E

5010 gag tgc ggt act tgg ttt aat acc cgt tct tgg aat gat aag gaa

35

! 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118  
 ! R Q P I I D W F L H A R K L G

5055 aga cag ccg att att gat tgg ttt cta cat gct cgt aaa tta gga

!

! 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133  
 ! W D I I F L V Q D L S I V D K  
 5100 tgg gat att att ttt ctt gtt cag gac tta tct att gtt gat aaa  
 !  
 5 ! 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148  
 ! Q A R S A L A E H V V Y C R R  
 5145 cag gcg cgt tct gca tta gct gaa cat gtt gtt tat tgt cgt cgt  
 !  
 10 ! 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163  
 ! L D R I T L P F V G T L Y S L  
 5190 ctg gac aga att act tta cct ttt gtc ggt act tta tat tct ctt  
 !  
 15 ! 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178  
 ! I T G S K M P L P K L H V G V  
 5235 att act ggc tgc aaa atg cct ctg cct aaa tta cat gtt ggc gtt  
 !  
 20 ! 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193  
 ! V K Y G D S Q L S P T V E R W  
 5280 gtt aaa tat ggc gat tct caa tta agc cct act gtt gag cgt tgg  
 !  
 25 ! 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208  
 ! L Y T G K N L Y N A Y D T K Q  
 5325 ctt tat act ggt aag aat ttg tat aac gca tat gat act aaa cag  
 !  
 30 ! 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223  
 ! A F S S N Y D S G V Y S Y L T  
 5370 gct ttt tct agt aat tat gat tcc ggt gtt tat tct tat tta acg  
 !  
 35 ! 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238  
 ! P Y L S H G R Y F K P L N L G  
 5415 cct tat tta tca cac ggt cgg tat ttc aaa cca tta aat tta ggt  
 !  
 ! 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253  
 ! Q K M K L T K I Y L K K F S R  
 5460 cag aag atg aaa tta act aaa ata tat ttg aaa aag ttt tct cgc  
 !  
 ! 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268  
 ! V L C L A I G F A S A F T Y S  
 5505 gtt ctt tgt ctt gcg att gga ttt gca tca gca ttt aca tat agt

119 120 121 122 123 124 125 126 127 128 129 130 131 132 133  
 W D I I F L V Q D L S I V D K  
 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148  
 Q A R S A L A E H V V Y C R R  
 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163  
 L D R I T L P F V G T L Y S L  
 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178  
 I T G S K M P L P K L H V G V  
 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193  
 V K Y G D S Q L S P T V E R W  
 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208  
 L Y T G K N L Y N A Y D T K Q  
 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223  
 A F S S N Y D S G V Y S Y L T  
 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238  
 P Y L S H G R Y F K P L N L G  
 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253  
 Q K M K L T K I Y L K K F S R  
 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268  
 V L C L A I G F A S A F T Y S

```

!      269 270 271 272 273 274 275 276 277 278 279 280 281 282 283
!      Y   I   T   Q   P   K   P   E   V   K   K   V   V   S   Q
5550 tat ata acc caa cct aag ccg gag gtt aaa aag gta gtc tct cag
5 !
!      284 285 286 287 288 289 290 291 292 293 294 295 296 297 298
!      T   Y   D   F   D   K   F   T   I   D   S   S   Q   R   L _
5595 acc tat gat ttt gat aaa ttc act att gac tct tct cag cgt ctt
!
10 !      299 300 301 302 303 304 305 306 307 308 309 310 311 312 313
!      N   L   S   Y   R   Y   V   F   K   D   S   K   G   K   L
5640 aat cta agc tat cgc tat gtt ttc aag gat tct aag gga aaa TTA
!                                     PacI
!
15 !      314 315 316 317 318 319 320 321 322 323 324 325 326 327 328
!      I   N   S   D   D   L   Q   K   Q   G   Y   S   L   T   Y
5685 ATT AAt agc gac gat tta cag aag caa ggt tat tca ctC aca tat
!      PacI
!
20 !      329 330 331 332 333 334 335 336 337 338 339 340 341 342 343
!      i   I   D   L   C   T   V   S   I   K   K   G   N   S   N   E
!      iv                                     M1 K
5730 att gat tta tgt act gtt tcc att aaa aaa ggt aat tca aAT Gaa
!                                     Start IV
25 !
!      344 345 346 347 348 349
!      i   I   V   K   C   N   .End of I
!      iv    L3 L   N5 V   I7 N   F V10
5775 att gtt aaa tgt aat TAA T TTT GTT
30 !      IV continued.....
5800 ttc ttg atg ttt gtt tca tca tct tct ttt gct cag gta att gaa atg
5848 aat aat tcg cct ctg cgc gat ttt gta act tgg tat tca aag caa tca
5896 ggc gaa tcc gtt att gtt tct ccc gat gta aaa ggt act gtt act gta
5944 tat tca tct gac gtt aaa cct gaa aat cta cgc aat ttc ttt att tct
35 5992 gtt tta cgt gct aat aat ttt gat atg gtt ggt tca att cct tcc ata
6040 att cag aag tat aat cca aac aat cag gat tat att gat gaa ttg cca
6088 tca tct gat aat cag gaa tat gat gat aat tcc gct cct tct ggt ggt
6136 ttc ttt gtt ccg caa aat gat aat gtt act caa act ttt aaa att aat
6184 aac gtt cgg gca aag gat tta ata cga gtt gtc gaa ttg ttt gta aag

```



6232 tct aat act tct aaa tcc tca aat gta tta tct att gac ggc tct aat  
 6280 cta tta gtt gtt TCT gca cct aaa gat att tta gat aac ctt cct caa  
 !  
 ApaLI removed  
 5 6328 ttc ctt tct act gtt gat ttg cca act gac cag ata ttg att gag ggt  
 6376 ttg ata ttt gag gtt cag caa ggt gat gct tta gat ttt tca ttt gct  
 6424 gct ggc tct cag cgt ggc act gtt gca ggc ggt gtt aat act gac cgc  
 6472 ctc acc tct gtt tta tct tct gct ggt ggt tct ggt att ttt\_aat  
 6520 ggc gat gtt tta ggg cta tca gtt cgc gca tta aag act aat agc cat  
 6568 tca aaa ata ttg tct gtg cca cgt att ctt acg ctt tca ggt cag aag  
 10 6616 ggt tct atc tct gtT GGC CAg aat gtc cct ttt att act ggt cgt gtg  
 !  
 MscI\_\_\_\_  
 6664 act ggt gaa tct gcc aat gta aat aat cca ttt cag acg att gag cgt  
 6712 caa aat gta ggt att tcc atg agc gtt ttt cct gtt gca atg gct ggc  
 6760 ggt aat att gtt ctg gat att acc agc aag gcc gat agt ttg agt tct  
 15 6808 tct act cag gca agt gat gtt att act aat caa aga agt att gct aca  
 6856 acg gtt aat ttg cgt gat gga cag act ctt tta ctc ggt ggc ctc act  
 6904 gat tat aaa aac act tct caa gat tct ggc gta ccg ttc ctg tct aaa  
 6952 atc cct tta atc ggc ctc ctg ttt agc tcc cgc tct gat tcc aac gag  
 7000 gaa agc acg tta tac gtg ctc gtc aaa gca acc ata gta cgc gcc ctg  
 20 7048 TAG cggcgcatt  
 !  
 End IV  
 7060 aagcgcggcg ggtgtggtgg ttacgcgcag cgtgaccgct acacttgcca gcgccttagc  
 7120 gcccgctcct ttcgctttct tcccttcctt tctcgccacg ttcGCCGGCt ttcccgctca  
 !  
 NgoMI\_  
 25 7180 agctctaaat cgggggctcc ctttaggggtt ccgatttagt gctttacggc acctcgacct  
 7240 caaaaaactt gatttgggtg atgggtCACG TAGTGggcca tcgcctgat agacgggttt  
 !  
 DraIII\_\_\_\_  
 7300 tcgccttttG ACGTTGGAGT Ccaggttctt taatagtga ctcttggtcc aaactggaac  
 !  
 DrdI\_\_\_\_  
 30 7360 aacactcaac cctatctcgg gctattcttt tgatttataa gggattttgc cgatttcgga  
 7420 accaccatca aacaggattt tcgcctgctg gggcaaacca gcgtggaccg cttgctgcaa  
 7480 ctctctcagg gccaggcggg gaagggaat CAGCTGttgc cCGTCTCact ggtgaaaaga  
 !  
 PvuII. BsmBI.  
 7540 aaaaccaccc tGGATCC AAGCTT  
 35 ! BamHI HindIII (1/2)  
 ! Insert carrying bla gene  
 7563 gcagggtg gcacttttctg gggaaatgtg cgcggaaccc  
 7600 ctatttggtt atttttctaa atacattcaa atatGTATCC gctcatgaga caataacct  
 !  
 BciVI

```

7660 gataaatgct tcaataatat tgaaaaAGGA AGAgt
!
! RBS.?...
!
! Start bla gene
5 7695 ATG agt att caa cat ttc cgt gtc gcc ctt att ccc ttt ttt gcg gca ttt
7746 tgc ctt cct gtt ttt gct cac cca gaa acg ctg gtg aaa gta aaa gat gct
7797 gaa gat cag ttg ggC gCA CGA Gtg ggt tac atc gaa ctg gat ctc aac agc
!
! BssSI...
!
! ApaLI removed
10 7848 ggt aag atc ctt gag agt ttt cgc ccc gaa gaa cgt ttt cca atg atg agc
7899 act ttt aaa gtt ctg cta tgt cat aca cta tta tcc cgt att gac gcc ggg
7950 caa gaG CAA CTC GGT CGc cgg gcg cgg tat tct cag aat gac ttg gtt gAG
!
! BcgI_____ ScaI
1001 TAC Tca cca gtc aca gaa aag cat ctt acg gat ggc atg aca gta aga gaa
!
! ScaI_
15 8052 tta tgc agt gct gcc ata acc atg agt gat aac act gcg gcc aac tta ctt
8103 ctg aca aCG ATC Gga gga ccg aag gag cta acc gct ttt ttg cac aac atg
!
! PvuI_____
18154 ggg gat cat gta act cgc ctt gat cgt tgg gaa ccg gag ctg aat gaa gcc
8205 ata cca aac gac gag cgt gac acc acg atg cct gta gca atg cca aca acg
20 8256 tTG CGC Aaa cta tta act ggc gaa cta ctt act cta gct tcc cgg caa caa
!
! FspI....
!
8307 tta ata gac tgg atg gag gcg gat aaa gtt gca gga cca ctt ctg cgc tcg
8358 GCC ctt ccG Gct ggc tgg ttt att gct gat aaa tct gga gcc ggt gag cgt
25 !
! BglI_____
8409 gGG TCT Cgc ggt atc att gca gca ctg ggg cca gat ggt aag ccc tcc cgt
!
! BsaI_____
8460 atc gta gtt atc tac acG ACg ggg aGT Cag gca act atg gat gaa cga aat
!
! AhdI_____
30 8511 aga cag atc gct gag ata ggt gcc tca ctg att aag cat tgg TAA ctgt
!
! stop
8560 cagaccaagt ttactcatat atactttaga ttgatttaaa acttcatttt taattttaaaa
8620 ggatctaggt gaagatcctt ttgataatc tcatgaccaa aatcccttaa cgtgagtttt
8680 cgttccactg tacgtaagac cccc
35 8704 AAGCTT GTCGAC tgaa tggcgaatgg cgctttgcct
!
! HindIII Sali..
!
! (2/2) HincII
8740 ggtttccggc accagaagcg gtgccggaaa gctggctgga gtgcgatctt
!

```

8790 CCTGAGG

! Bsu36I\_

8797 ccgat actgtcgtcg tccoctcaaa ctggcagatg

8832 cacgggttacg atgcgcccac ctacaccaac gtaacctatc ccattacggt caatccgccg

5 8892 tttgttccca cggagaatcc gacgggttgt tactcgctca catttaatgt tgatgaaagc

8952 tggctacagg aaggccagac gcgaattatt tttgatggcg ttcctattgg ttaaaaaatg

9012 agctgattta acaaaaaattt aacgcgaatt ttaacaaaat attaacgttt acaATTAAAA

! SwaI...

9072 Tatttgctta tacaatcttc ctgtttttgg ggcttttctg attatcaacc GGGGTACat

10 ! RBS?

9131 ATG att gac atg cta gtt tta cga tta ccg ttc atc gat tct ctt gtt tgc

! Start gene II

9182 tcc aga ctc tca ggc aat gac ctg ata gcc ttt gtA GAT CTc tca aaa ata

! BglIII...

15 9233 gct acc ctc tcc ggc atg aat tta tca gct aga acg gtt gaa tat cat att

9284 gat ggt gat ttg act gtc tcc ggc ctt tct cac cct ttt gaa tct tta cct

9335 aca cat tac tca ggc att gca ttt aaa ata tat gag ggt tct aaa aat ttt

9386 tat cct tgc gtt gaa ata aag gct tct ccc gca aaa gta tta cag ggt cat

9437 aat gtt ttt ggt aca acc gat tta gct tta tgc tct gag gct tta ttg ctt

20 9488 aat ttt gct aat tct ttg cct tgc ctg tat gat tta ttg gat gtt ! 9532

! gene II continues

Table 120B: Sequence of MALIA3, condensed

LOCUS	MALIA3	9532	CIRCULAR
ORIGIN			
1	AATGCTACTA	CTATTAGTAG	AATTGATGCC
5	61 ATAGCTAAAC	AGGTTATTGA	CCATTGCGA
121	CGTTCGCAGA	ATTGGGAATC	AACTGTTACA
181	GTTGCATATT	TAAAACATGT	TGAGCTACAG
241	TCCGCAAAAA	TGACCTCTTA	TCAAAAGGAG
301	TTGGAGTTTG	CTTCCGGTCT	GGTTCGCTTT
10	361 TCTTTCGGGC	TTCCTCTTAA	TCTTTTGGAT
421	CAGGGTAAAG	ACCTGATTTT	TGATTTATGG
481	TTTGAGGGGG	ATTCAATGAA	TATTTATGAC
541	AAACATTTTA	CTATTACCCC	CTCTGGCAAA
601	GGTTTTTATC	GTCTCTGGT	AAACGAGGGT
15	661 AATTCCTTTT	GGCGTTATGT	ATCTGCATTA
721	ATGAATCTTT	CTACCTGTAA	TAATGTTGTT
781	TCTTCCCAAC	GTCTTGACTG	GTATAATGAG
841	CAATGATTAA	AGTTGAAATT	AAACCATCTC
901	CTCGTCAGGG	CAAGCCTTAT	TCACTGAATG
20	961 AATATCCGGT	TCTTGTC AAG	ATTACTCTTG
1021	TGTACACCGT	TCATCTGTCC	TCTTTC AAG
1081	GTCTGCGCCT	CGTTCGGGCT	AAGTAACATG
1141	CAGGCGATGA	TACAAATCTC	CGTTGTACTT
1201	CAAAGATGAG	TGTTTTAGTG	TATCTTTTCG
25	1261 GTGGCATTAC	GTATTTTACC	CGTTAATG
1321	CAAAGCCTCT	GAGCCGTTG	CTACCTCGT
1381	CGATCCCGCA	AAAGCGGCCT	TTAACTCCCT
1441	TGCGTGGGCG	ATGGTTGTTG	TCATTGTGCG
1501	ATTACCTCG	AAAGCAAGCT	GATAAACCGA
30	1561 TTTTGGGAGA	TTTTC AACGT	GAAAAAATTA
1621	TATTCTCACA	GTGCACAGTC	TGTCGTGACG
1681	CAGAGGGTCA	CCATCTCCTG	CACTGGGAGC
1741	CACTGGTACC	AGCAGCTTCC	AGGAACAGCC
1801	CGGCCCTCAG	GGGTCCCTGA	CCGATTCTCT
35	1861 GCCATCACTG	GGCTCCAGGC	TGAGGATGAG
1921	AGCCTGAGTG	GCCTTTATGT	CTTCGGA
1981	AAGGCCAACC	CCACTGTCAC	TCTGTTCCCG
2041	GCCACACTAG	TGTGTCTGAT	CAGTGACTTC
2101	GCAGATAGCA	GCCCCGTCAA	GGCGGGAGTG

5

10

15

20

25

30

35

2161 AACAAGTACG CGGCCAGCAG CTATCTGAGC CTGACGCCTG AGCAGTGGAA GTCCACACAGA  
2221 AGCTACAGCT GCCAGGTCAC GCATGAAGGG AGCACCCTGG AGAAGACAGT GGCCCTTACA  
2281 GAATGTTTCAT AATAAACCGC CTCCACCGGG CGCGCCAATT CTATTTCAAG GAGACAGTCA  
2341 TAATGAAATA CCTATTGCCT ACGGCAGCCG CTGGATTGTT ATTACTCGCG GCCCAGCCGG  
2401 CCATGGCCGA AGTTCAATTG TTAGAGTCTG GTGGCGGTCT TGTTCAGCCT GGTGGTTCTT  
2461 TACGTCTTTC TTGCGCTGCT TCCGGATTCA CTTTCTCTTC GTACGCTATG TCTTGGGTTC  
2521 GCCAAGCTCC TGGTAAAGGT TTGGAGTGGG TTTCTGCTAT CTCTGGTTCT GGTGGCAGTA  
2581 CTACTATGC TGA CTCCGTT AAAGGTCGCT TCACTATCTC TAGAGACAAC TCTAAGAATA  
2641 CTCTCTACTT GCAGATGAAC AGCTTAAGGG CTGAGGACAC TGCAGTCTAC TATTGCGCTA  
2701 AAGACTATGA AGGTACTGGT TATGCTTTCG ACATATGGGG TCAAGGTACT ATGGTCACCG  
2761 TCTCTAGTGC CTCCACCAAG GGCCCATCGG TCTTCCCCCT GGCACCCTCC TCCAAGAGCA  
2821 CCTCTGGGGG CACAGCGGCC CTGGGCTGCC TGGTCAAGGA CTACTTCCCC GAACCGGTGA  
2881 CGGTGTCGTG GAACTCAGGC GCCCTGACCA GCGGCGTCCA CACCTTCCCC GCTGTCTTAC  
2941 AGTCTAGCGG ACTCTACTCC CTCAGCAGCG TAGTGACCGT GCCCTCTTCT AGCTTGGGCA  
3001 CCCAGACCTA CATCTGCAAC GTGAATCACA AGCCCAGCAA CACCAAGGTG GACAAGAAAG  
3061 TTGAGCCCAA ATCTTGTCG GCGGCTCATC ACCACCATCA TCACTCTGCT GAACAAAAAC  
3121 TCATCTCAGA AGAGGATCTG AATGGTGCCG CAGATATCAA CGATGATCGT ATGGCTGGCG  
3181 CCGCTGAAAC TGTGAAAGT TGTTTAGCAA AACCCCATAC AGAAAATTCA TTTACTAACG  
3241 TCTGGAAAGA CGACAAAAC TTAGATCGTT ACGCTAACTA TGAGGGTTGT CTGTGGAATG  
3301 CTACAGGCGT TGTAGTTTGT ACTGGTGACG AAATCAGTG TTACGGTACA TGGGTTTCTA  
3361 TTGGGCTTGC TATCCCTGAA AATGAGGGTG GTGGCTCTGA GGGTGGCGGT TCTGAGGGTG  
3421 GCGTTTCTGA GGGTGGCGGT ACTAAACCTC CTGAGTACGG TGATACACCT ATTCCGGGCT  
3481 ATACTTATAT CAACCTCTC GACGGCACTT ATCCGCCTGG TACTGAGCAA AACCCGCTA  
3541 ATCCTAATCC TTCTCTTGG GAGTCTCAGC CTCTTAATAC TTTTATGTTT CAGAATAATA  
3601 GGTTCGAAA TAGGCAGGGG GCATTAAC TGTTTATACGG CACTGTTACT CAAGGCACTG  
3661 ACCCGTTAA AACTTATTAC CAGTACACTC CTGTATCATC AAAAGCCATG TATGACGCTT  
3721 ACTGGAACGG TAAATTCAGA GACTGCGCTT TCCATTCTGG CTTTAATGAA GATCCATTG  
3781 TTTGTGAATA TCAAGGCCAA TCGTCTGACC TGCCTCAACC TCCTGTCAAT GCTGGCGGCG  
3841 GCTCTGGTGG TGGTCTGGT GCGGCTCTG AGGGTGGTGG CTCTGAGGGT GCGGTTCTG  
3901 AGGGTGGCGG CTCTGAGGGA GCGGTTCCG GTGGTGGCTC TGGTTCGGT GATTTTGATT  
3961 ATGAAAAGAT GGCAAACGCT AATAAGGGGG CTATGACCGA AAATGCCGAT GAAAACGCGC  
4021 TACAGTCTGA CGCTAAAGGC AAAC TTGATT CTGTCGCTAC TGATTACGGT GCTGCTATCG  
4081 ATGGTTTCAT TGGTGACGTT TCCGGCCTTG CTAATGGTAA TGGTGCTACT GGTGATTTTG  
4141 CTGGCTCTAA TTCCCAAATG GCTCAAGTCG GTGACGGTGA TAATTCACCT TTAATGAATA  
4201 ATTTCCGTCA ATATTTACCT TCCCTCCCTC AATCGGTTGA ATGTCGCCCT TTTGTCTTTA  
4261 GCGCTGGTAA ACCATATGAA TTTTCTATTG ATTTGTGACAA AATAAACTTA TTCCGTGGTG  
4321 TCTTTGCGTT TCTTTTATAT GTTGCCACCT TTATGTATGT ATTTTCTACG TTTGCTAACA  
4381 TACTGCGTAA TAAGGAGTCT TAATCATGCC AGTTCTTTTG GGTATTCCGT TATTATTGCG  
4441 TTTCTCGGT TTCCTTCTGG TAACTTTGTT CGGCTATCTG CTTACTTTTC TAAAAAGGG

5  
 10  
 15  
 20  
 25  
 30  
 35

4501 CTTTCGGTAAG ATAGCTATTG CTATTTTCATT GTTTCCTTGCT CTTATTATTG GGCTTAACTC  
 4561 AATTCTTGTG GGTTATCTCT CTGATATTAG CGCTCAATTA CCCTCTGACT TTGTTTCAGGG  
 4621 TGTTTCAGTTA ATTCTCCCGT CTAATGCGCT TCCCTGTTTT TATGTTATTCT TCTCTGTAAA  
 4681 GGCTGCTATT TTCATTTTTG ACGTTAAACA AAAAATCGTT TCTTATTTGG ATTGGGATAA  
 4741 ATAATATGGC TGTTTATTTT GTAAGTGGCA AATTAGGCTC TGGAAAGACG CTCGTTAGCG  
 4801 TTGGTAAGAT TCAGGATAAA ATTGTAGCTG GGTGCAAAAT AGCAACTAAT CTTGATTTAA  
 4861 GGCTTCAAAA CCTCCCGCAA GTCGGGAGGT TCGCTAAAAC GCCTCGCGTT CTAGAAATAC  
 4921 CGGATAAGCC TTCTATATCT GATTTGCTTG CTATTGGGCG CGGTAATGAT TCCTACGATG  
 4981 AAAATAAAAA CGGCTTGCTT GTTCTCGATG AGTGCGGTAC TTGGTTTAAAT ACCCGTTCTT  
 5041 GGAATGATAA GGAAAGACAG CCGATTATTG ATTGGTTTCT ACATGCTCGT AAATTAGGAT  
 5101 GGGATATTAT TTTTCTTGTT CAGGACTTAT CTATTGTTGA TAAACAGGCG CGTTCTGCAT  
 5161 TAGCTGAACA TGTTGTTTAT TGTCGTCGTC TGGACAGAAT TACTTTACCT TTTGTCGGTA  
 5221 CTTTATATTCT TCTTATTACT GGCTCGAAAA TGCCTCTGCC TAAATTACAT GTTGGCGTTG  
 5281 TTAAATATGG CGATTCTCAA TTAAGCCCTA CTGTTGAGCG TTGGCTTTAT ACTGGTAAGA  
 5341 ATTTGTATAA CGCATATGAT ACTAAACAGG CTTTTTCTAG TAATTATGAT TCCGGTGTTT  
 5401 ATCTTATTTT AACGCCCTAT TTATCACACG GTCGGTATTT CAAACCATTAA AATTTAGGTC  
 5461 AGAAGATGAA ATTAACATAA ATATATTTGA AAAAGTTTTT TCGCGTTCTT TGTCTTGCGA  
 5521 TTGGATTGTC ATCAGCATTT ACATATAGTT ATATAACCCA ACCTAAGCCG GAGGTTAAAA  
 5581 AGGTAGTCTC TCAGACCTAT GATTTTGATA AATTCACAT TGAATCTTCT CAGCGTCTTA  
 5641 ATCTAAGCTA TCGCTATGTT TTCAAGGATT CTAAGGGAAA ATTAATTAAT AGCGACGATT  
 5701 TACAGAAGCA AGGTTATTCA CTCACATATA TTGATTTATG TACTGTTTCC ATTAATAAAG  
 5761 GTAATTCAAA TGAAATGTTT AAATGTAATT AATTTTGTTT TCTTGATGTT TGTTTCATCA  
 5821 TCTTCTTTTG CTCAGGTAAT TGAAATGAAT AATTCGCCTC TGCGCGATTT TGTAAGTTGG  
 5881 TATTCAAAGC AATCAGGCGA ATCCGTTATT GTTCTCCCG ATGTAAAAGG TACTGTTACT  
 5941 GTATATTCAT CTGACGTTAA ACCTGAAAAT CTACGCAATT TCTTTATTTT TGTTTTACGT  
 6001 GCTAATAATT TTGATATGGT TGGTTCAATT CCTTCCATAA TTCAGAAGTA TAATCCAAAC  
 6061 AATCAGGATT ATATTGATGA ATTGCCATCA TCTGATAATC AGGAATATGA TGATAATTCC  
 6121 GCTCCTTCTG GTGGTTTCTT TGTTCCGCAA AATGATAATG TTAATCAAAC TTTTAAAATT  
 6181 AATAACGTTT GGGCAAAGGA TTTAATACGA GTTGTGCAAT TGTTTGTAAG GTCTAATACT  
 6241 TCTAAATCCT CAAATGTATT ATCTATTGAC GGCTCTAATC TATTAGTTGT TTCTGCACCT  
 6301 AAAGATATTT TAGATAACCT TCCTCAATTC CTTTCTACTG TTGATTTGCC AACTGACCAG  
 6361 ATATTGATTG AGGGTTTGAT ATTTGAGGTT CAGCAAGGTG ATGCTTTAGA TTTTTCATTT  
 6421 GCTGCTGGCT CTCAGCGTGG CACTGTTGCA GCGGGTGTTA ATACTGACCG CCTCACCTCT  
 6481 GTTTTATCTT CTGCTGGTGG TTCGTTCCGGT ATTTTAAATG GCGATGTTTT AGGGCTATCA  
 6541 GTTCGCGCAT TAAAGACTAA TAGCCATTCA AAAATATTGT CTGTGCCACG TATTCTTACG  
 6601 CTTTCAGGTC AGAAGGGTTC TATCTCTGTT GGCCAGAATG TCCCTTTTAT TACTGGTCGT  
 6661 GTGACTGGTG AATCTGCCAA TGTAATAAAT CCATTTTACA CGATTGAGCG TCAAAATGTA  
 6721 GGTATTTCCA TGAGCGTTTT TCCTGTTGCA ATGGCTGGCG GTAATATTGT TCTGGATATT  
 6781 ACCAGCAAGG CCGATAGTTT GAGTCTTCT ACTCAGGCAA GTGATGTTAT TACTAATCAA

5  
10  
15  
20  
25  
30  
35

6841 AGAAGTATTG CTACAACGGT TAATTTGCGT GATGGACAGA CTCTTTTACT CGGTGGCCTC  
6901 ACTGATTATA AAAACACTTC TCAAGATTCT GGCGTACCGT TCCTGTCTAA AATCCCTTTA  
6961 ATCGGCCTCC TGTTTAGCTC CCGCTCTGAT TCCAACGAGG AAAGCACGTT ATACGTGCTC  
7021 GTCAAAGCAA CCATAGTACG CGCCCTGTAG CGGCGCATT AAGCGGCGG GTGTGGTGGT  
7081 TACGCGCAGC GTGACCGCTA CACTTGCCAG CGCCCTAGCG CCCGCTCCTT TCGCTTTCTT  
7141 CCCTTCCTTT CTCGCCACGT TCGCCGGCTT TCCCGTCAA GCTCTAAATC GGGGGCTCCC  
7201 TTTAGGGTTC CGATTTAGTG CTTTACGGCA CCTCGACCCC AAAAACTTG ATTGGGTGA  
7261 TGGTTCACGT AGTGGGCCAT CGCCCTGATA GACGGTTTTT CGCCCTTTGA CGTTGGAGTC  
7321 CACGTTCTTT AATAGTGGAC TCTTGTTCCA AACTGGAACA ACACTCAACC CTATCTCGGG  
7381 CTATTCTTTT GATTTATAAG GGATTTTGCC GATTTTCGAA CCACCATCAA ACAGGATTTT  
7441 CGCCTGCTGG GGCAAACCAG CGTGGACCGC TTGCTGCAAC TCTCTCAGGG CCAGGCGGTG  
7501 AAGGGCAATC AGCTGTTGCC CGTCTCACTG GTGAAAAGAA AAACCACCCT GGATCCAAGC  
7561 TTGCAGGTGG CACTTTTCGG GGAAATGTGC GCGGAACCCC TATTTGTTTA TTTTCTAAA  
7621 TACATTCAAA TATGTATCCG CTCATGAGAC AATAACCTG ATAAATGCTT CAATAATATT  
7681 GAAAAAGGAA GAGTATGAGT ATTCAACATT TCCGTGTCGC CCTTATTCCC TTTTTGCGG  
7741 CATTTTGCCT TCCTGTTTTT GTCACCCAG AAACGCTGGT GAAAGTAAAA GATGCTGAAG  
7801 ATCAGTTGGG CGCACGAGTG GGTACATCG AACTGGATCT CAACAGCGGT AAGATCCTTG  
7861 AGAGTTTTCG CCCCAGAGAA CGTTTTCCAA TGATGAGCAC TTTTAAAGTT CTGCTATGTC  
7921 ATACACTATT ATCCGTATT GACGCCGGG AAGAGCAACT CGGTGCGCGG GCGCGGTATT  
7981 CTCAGAATGA CTGGTTGAG TACTCACCAG TCACAGAAAA GCATCTTACG GATGGCATGA  
8041 CAGTAAGAGA ATTAGCAGT GCTGCCATAA CCATGAGTGA TAACACTGCG GCCAACTTAC  
8101 TTCTGACAAC GATCGGAGGA CCGAAGGAGC TAACCGCTT TTTGCACAAC ATGGGGGATC  
8161 ATGTAACCTG CCTTGATCGT TGGGAACCGG AGCTGAATGA AGCCATACCA AACGACGAGC  
8221 GTGACACCAC GATGCCGTGA GCAATGCCAA CAACGTTGCG CAAACTATTA ACTGGCGAAC  
8281 TACTTACTCT AGCTTCCCG CAACAATTAA TAGACTGGAT GGAGGCGGAT AAAGTTGCAG  
8341 GACCACTTCT GCGCTCGGCC CTTCGGCTG GCTGGTTTAT TGCTGATAAA TCTGGAGCCG  
8401 GTGAGCGTGG GTCTCGCGGT ATCATTGAG CACTGGGGCC AGATGGTAAG CCTCCCCTA  
8461 TCGTAGTTAT CTACACGACG GGGAGTCAGG CAACTATGGA TGAACGAAAT AGACAGATCG  
8521 CTGAGATAGG TGCTCACTG ATTAAGCATT GGTAACGTG AGACCAAGTT TACTCATATA  
8581 TACTTTAGAT TGATTTAAAA CTTCATTTT AATTTAAAAG GATCTAGGTG AAGATCCTTT  
8641 TTGATAATCT CATGACCAAA ATCCCTTAAC GTGAGTTTTC GTTCCACTGT ACGTAAGACC  
8701 CCAAGCTTG TCGACTGAAT GGCGAATGGC GCTTGCCTG GTTCCGGCA CCAGAAGCGG  
8761 TGCCGGAAG CTGGCTGGAG TCGATCTTC CTGAGGCCGA TACTGTCGTC GTCCCTCAA  
8821 ACTGGCAGAT GCACGGTTAC GATGCGCCCA TCTACACCAA CGTAACCTAT CCCATTACGG  
8881 TCAATCCGCC GTTGTTCCT ACAGGAGATC CGACGGGTTG TTACTCGCTC ACATTTAATG  
8941 TTGATGAAAG CTGGCTACAG GAAGGCCAGA CGCGAATTAT TTTTGATGGC GTTCCTATTG  
9001 GTTAAAAAAT GAGCTGATTT AACAAAAATT TAACGCGAAT TTTAACAAAA TATTAACGTT  
9061 TACAATTTAA ATATTGCTT ATACAATCTT CCTGTTTTT GGGCTTTTCT GATTATCAAC  
9121 CGGGGTACAT ATGATTGACA TGCTAGTTTT ACGATTACCG TTCATCGATT CTCTTGTTTG

9181 CTCCAGACTC TCAGGCAATG ACCTGATAGC CTTTGTAGAT CTCTCAAAAA TAGCTACCCT  
9241 CTCCGGCATG AATTTATCAG CTAGAACGGT TGAATATCAT ATTGATGGTG ATTTGACTGT  
9301 CTCCGGCCTT TCTCACCTT TTGAATCTTT ACCTACACAT TACTCAGGCA TTGCATTTAA  
9361 AATATATGAG GGTTCATAAA ATTTTATCC TTGCGTTGAA ATAAAGGCTT CTCCCGCAA  
9421 AGTATTACAG GGTCATAATG TTTTGGTAC AACCGATTTA GCTTTATGCT CTGAGGCTTT  
9481 ATTGCTTAAT TTTGCTAATT CTTTGCCTTG CCTGTATGAT TTATTGGATG TT



Table 200: Enzymes that either cut 15 or more human GLGs or have 5+-base recognition in FR3

Typical entry:

RName Recognition #sites  
 GLGid#:base# GLGid#:base# GLGid#:base#.....

5

BstEII Ggtnacc 2  
 1: 3 48: 3  
 There are 2 hits at base# 3

10

MaeIII gtnac 36  
 1: 4 2: 4 3: 4 4: 4 5: 4 6: 4  
 7: 4 8: 4 9: 4 10: 4 11: 4 37: 4  
 37: 58 38: 4 38: 58 39: 4 39: 58 40: 4  
 40: 58 41: 4 41: 58 42: 4 42: 58 43: 4  
 43: 58 44: 4 44: 58 45: 4 45: 58 46: 4  
 46: 58 47: 4 47: 58 48: 4 49: 4 50: 58  
 There are 24 hits at base# 4

15

Tsp45I gtsac 33  
 1: 4 2: 4 3: 4 4: 4 5: 4 6: 4  
 7: 4 8: 4 9: 4 10: 4 11: 4 37: 4  
 37: 58 38: 4 38: 58 39: 58 40: 4 40: 58  
 41: 58 42: 58 43: 4 43: 58 44: 4 44: 58  
 45: 4 45: 58 46: 4 46: 58 47: 4 47: 58  
 48: 4 49: 4 50: 58  
 There are 21 hits at base# 4

20

25

HphI tcacc 45  
 1: 5 2: 5 3: 5 4: 5 5: 5 6: 5  
 7: 5 8: 5 11: 5 12: 5 12: 11 13: 5  
 14: 5 15: 5 16: 5 17: 5 18: 5 19: 5  
 20: 5 21: 5 22: 5 23: 5 24: 5 25: 5  
 26: 5 27: 5 28: 5 29: 5 30: 5 31: 5  
 32: 5 33: 5 34: 5 35: 5 36: 5 37: 5  
 38: 5 40: 5 43: 5 44: 5 45: 5 46: 5  
 47: 5 48: 5 49: 5  
 There are 44 hits at base# 5

30

35

NlaIII CATG

26

1: 9 1: 42 2: 42 3: 9 3: 42 4: 9  
4: 42 5: 9 5: 42 6: 42 6: 78 7: 9  
7: 42 8: 21 8: 42 9: 42 10: 42 11: 42  
5 12: 57 13: 48 13: 57 14: 57 31: 72 38: 9  
48: 78 49: 78

There are 11 hits at base# 42

There are 1 hits at base# 48 Could cause raggedness.

10 BsaJI Ccnngg

37

1: 14 2: 14 5: 14 6: 14 7: 14 8: 14  
8: 65 9: 14 10: 14 11: 14 12: 14 13: 14  
14: 14 15: 65 17: 14 17: 65 18: 65 19: 65  
20: 65 21: 65 22: 65 26: 65 29: 65 30: 65  
15 33: 65 34: 65 35: 65 37: 65 38: 65 39: 65  
40: 65 42: 65 43: 65 48: 65 49: 65 50: 65  
51: 14

There are 23 hits at base# 65

There are 14 hits at base# 14

20

AluI AGct

42

1: 47 2: 47 3: 47 4: 47 5: 47 6: 47  
7: 47 8: 47 9: 47 10: 47 11: 47 16: 63  
23: 63 24: 63 25: 63 31: 63 32: 63 36: 63  
25 37: 47 37: 52 38: 47 38: 52 39: 47 39: 52  
40: 47 40: 52 41: 47 41: 52 42: 47 42: 52  
43: 47 43: 52 44: 47 44: 52 45: 47 45: 52  
46: 47 46: 52 47: 47 47: 52 49: 15 50: 47

There are 23 hits at base# 47

30 There are 11 hits at base# 52 Only 5 bases from 47

BlpI GCtnagc

21

1: 48 2: 48 3: 48 5: 48 6: 48 7: 48  
8: 48 9: 48 10: 48 11: 48 37: 48 38: 48  
35 39: 48 40: 48 41: 48 42: 48 43: 48 44: 48  
45: 48 46: 48 47: 48

There are 21 hits at base# 48

MwoI GCNNNNNngc

19

1: 48 2: 28 19: 36 22: 36 23: 36 24: 36  
25: 36 26: 36 35: 36 37: 67 39: 67 40: 67  
41: 67 42: 67 43: 67 44: 67 45: 67 46: 67  
47: 67

5

There are 10 hits at base# 67

There are 7 hits at base# 36

DdeI Ctnag

71

10

1: 49 1: 58 2: 49 2: 58 3: 49 3: 58  
3: 65 4: 49 4: 58 5: 49 5: 58 5: 65  
6: 49 6: 58 6: 65 7: 49 7: 58 7: 65  
8: 49 8: 58 9: 49 9: 58 9: 65 10: 49  
10: 58 10: 65 11: 49 11: 58 11: 65 15: 58

15

16: 58 16: 65 17: 58 18: 58 20: 58 21: 58  
22: 58 23: 58 23: 65 24: 58 24: 65 25: 58  
25: 65 26: 58 27: 58 27: 65 28: 58 30: 58  
31: 58 31: 65 32: 58 32: 65 35: 58 36: 58  
36: 65 37: 49 38: 49 39: 26 39: 49 40: 49

20

41: 49 42: 26 42: 49 43: 49 44: 49 45: 49  
46: 49 47: 49 48: 12 49: 12 51: 65

There are 29 hits at base# 58

There are 22 hits at base# 49 Only nine base from 58

There are 16 hits at base# 65 Only seven bases from 58

25

BglII Agatct

11

1: 61 2: 61 3: 61 4: 61 5: 61 6: 61  
7: 61 9: 61 10: 61 11: 61 51: 47

There are 10 hits at base# 61

30

BstYI Rgatcy

12

1: 61 2: 61 3: 61 4: 61 5: 61 6: 61  
7: 61 8: 61 9: 61 10: 61 11: 61 51: 47

There are 11 hits at base# 61

35

Hpy188I TCNga 17  
 1: 64 2: 64 3: 64 4: 64 5: 64 6: 64  
 7: 64 8: 64 9: 64 10: 64 11: 64 16: 57  
 20: 57 27: 57 35: 57 48: 67 49: 67

5 There are 11 hits at base# 64  
 There are 4 hits at base# 57  
 There are 2 hits at base# 67 Could be ragged.

MslI CAYNNnnRTG 44  
 10 1: 72 2: 72 3: 72 4: 72 5: 72 6: 72  
 7: 72 8: 72 9: 72 10: 72 11: 72 15: 72  
 17: 72 18: 72 19: 72 21: 72 23: 72 24: 72  
 25: 72 26: 72 28: 72 29: 72 30: 72 31: 72  
 32: 72 33: 72 34: 72 35: 72 36: 72 37: 72  
 15 38: 72 39: 72 40: 72 41: 72 42: 72 43: 72  
 44: 72 45: 72 46: 72 47: 72 48: 72 49: 72  
 50: 72 51: 72  
 There are 44 hits at base# 72

20 BsiEI CGRYcg 23  
 1: 74 3: 74 4: 74 5: 74 7: 74 8: 74  
 9: 74 10: 74 11: 74 17: 74 22: 74 30: 74  
 33: 74 34: 74 37: 74 38: 74 39: 74 40: 74  
 41: 74 42: 74 45: 74 46: 74 47: 74  
 25 There are 23 hits at base# 74

EaeI Yggccr 23  
 1: 74 3: 74 4: 74 5: 74 7: 74 8: 74  
 9: 74 10: 74 11: 74 17: 74 22: 74 30: 74  
 30 33: 74 34: 74 37: 74 38: 74 39: 74 40: 74  
 41: 74 42: 74 45: 74 46: 74 47: 74  
 There are 23 hits at base# 74

EagI Cggccg 23  
 35 1: 74 3: 74 4: 74 5: 74 7: 74 8: 74  
 9: 74 10: 74 11: 74 17: 74 22: 74 30: 74

1. The first of these is the fact that the  
 2. of the first of these is the fact that the  
 3. of the first of these is the fact that the  
 4. of the first of these is the fact that the  
 5. of the first of these is the fact that the  
 6. of the first of these is the fact that the  
 7. of the first of these is the fact that the  
 8. of the first of these is the fact that the  
 9. of the first of these is the fact that the  
 10. of the first of these is the fact that the

5

27

10

Bst4CI ACNgt 65°C

**63 Sites** There is a third isoschismer

15

20

25

There are 51 hits at base# 86 All the other sites are well away

63

30

35

44: 86 45: 86 46: 86 47: 86 48: 86 49: 86  
 50: 86 51: 0 51: 86

There are 51 hits at base# 86

5 HinfI Gantc

43

2: 2 3: 2 4: 2 5: 2 6: 2 7: 2  
 8: 2 9: 2 9: 22 10: 2 11: 2 15: 2  
 16: 2 17: 2 18: 2 19: 2 19: 22 20: 2  
 21: 2 23: 2 24: 2 25: 2 26: 2 27: 2  
 28: 2 29: 2 30: 2 31: 2 32: 2 33: 2  
 33: 22 34: 22 35: 2 36: 2 37: 2 38: 2  
 40: 2 43: 2 44: 2 45: 2 46: 2 47: 2  
 50: 60

There are 38 hits at base# 2

15

MlyI GAGTCNNNNNn

18

2: 2 3: 2 4: 2 5: 2 6: 2 7: 2  
 8: 2 9: 2 10: 2 11: 2 37: 2 38: 2  
 40: 2 43: 2 44: 2 45: 2 46: 2 47: 2

20

There are 18 hits at base# 2

PleI gagtc

18

2: 2 3: 2 4: 2 5: 2 6: 2 7: 2  
 8: 2 9: 2 10: 2 11: 2 37: 2 38: 2  
 40: 2 43: 2 44: 2 45: 2 46: 2 47: 2

25

There are 18 hits at base# 2

AciI Ccgc

24

2: 26 9: 14 10: 14 11: 14 27: 74 37: 62  
 37: 65 38: 62 39: 65 40: 62 40: 65 41: 65  
 42: 65 43: 62 43: 65 44: 62 44: 65 45: 62  
 46: 62 47: 62 47: 65 48: 35 48: 74 49: 74

30

There are 8 hits at base# 62

There are 8 hits at base# 65

There are 3 hits at base# 14

35

There are 3 hits at base# 74

There are 1 hits at base# 26

There are 1 hits at base# 35

-- Gcgg 11  
 8: 91 9: 16 10: 16 11: 16 37: 67 39: 67  
 40: 67 42: 67 43: 67 45: 67 46: 67

There are 7 hits at base# 67

5 There are 3 hits at base# 16

There are 1 hits at base# 91

BsiHKAI GWGCWc 20  
 2: 30 4: 30 6: 30 7: 30 9: 30 10: 30  
 10 12: 89 13: 89 14: 89 37: 51 38: 51 39: 51  
 40: 51 41: 51 42: 51 43: 51 44: 51 45: 51  
 46: 51 47: 51

There are 11 hits at base# 51

15 Bsp1286I GDGCHc 20  
 2: 30 4: 30 6: 30 7: 30 9: 30 10: 30  
 12: 89 13: 89 14: 89 37: 51 38: 51 39: 51  
 40: 51 41: 51 42: 51 43: 51 44: 51 45: 51  
 46: 51 47: 51

20 There are 11 hits at base# 51

HgiAI GWGCWc 20  
 2: 30 4: 30 6: 30 7: 30 9: 30 10: 30  
 12: 89 13: 89 14: 89 37: 51 38: 51 39: 51  
 25 40: 51 41: 51 42: 51 43: 51 44: 51 45: 51  
 46: 51 47: 51

There are 11 hits at base# 51

BsoFI GCngc 26  
 30 2: 53 3: 53 5: 53 6: 53 7: 53 8: 53  
 8: 91 9: 53 10: 53 11: 53 31: 53 36: 36  
 37: 64 39: 64 40: 64 41: 64 42: 64 43: 64  
 44: 64 45: 64 46: 64 47: 64 48: 53 49: 53  
 50: 45 51: 53

35 There are 13 hits at base# 53

There are 10 hits at base# 64

TseI GcwgC 17  
 2: 53 3: 53 5: 53 6: 53 7: 53 8: 53

9: 53 10: 53 11: 53 31: 53 36: 36 45: 64  
46: 64 48: 53 49: 53 50: 45 51: 53

There are 13 hits at base# 53

5 MnlI gagg

34

3: 67 3: 95 4: 51 5: 16 5: 67 6: 67  
7: 67 8: 67 9: 67 10: 67 11: 67 15: 67  
16: 67 17: 67 19: 67 20: 67 21: 67 22: 67  
23: 67 24: 67 25: 67 26: 67 27: 67 28: 67  
10 29: 67 30: 67 31: 67 32: 67 33: 67 34: 67  
35: 67 36: 67 50: 67 51: 67

There are 31 hits at base# 67

HpyCH4V TGca

34

15 5: 90 6: 90 11: 90 12: 90 13: 90 14: 90  
15: 44 16: 44 16: 90 17: 44 18: 90 19: 44  
20: 44 21: 44 22: 44 23: 44 24: 44 25: 44  
26: 44 27: 44 27: 90 28: 44 29: 44 33: 44  
34: 44 35: 44 35: 90 36: 38 48: 44 49: 44  
20 50: 44 50: 90 51: 44 51: 52

There are 21 hits at base# 44

There are 1 hits at base# 52

AccI GTmkac

13 5-base recognition

25 7: 37 11: 24 37: 16 38: 16 39: 16 40: 16  
41: 16 42: 16 43: 16 44: 16 45: 16 46: 16  
47: 16

There are 11 hits at base# 16

30 SacII CCGCg

8 6-base recognition

9: 14 10: 14 11: 14 37: 65 39: 65 40: 65  
42: 65 43: 65

There are 5 hits at base# 65

There are 3 hits at base# 14

35

TfiI Gawtc

24

9: 22 15: 2 16: 2 17: 2 18: 2 19: 2  
19: 22 20: 2 21: 2 23: 2 24: 2 25: 2



26: 2 27: 2 28: 2 29: 2 30: 2 31: 2  
 32: 2 33: 2 33: 22 34: 22 35: 2 36: 2  
 There are 20 hits at base# 2

5 BsmAI Nnnnnngagac 19  
 15: 11 16: 11 20: 11 21: 11 22: 11 23: 11  
 24: 11 25: 11 26: 11 27: 11 28: 11 28: 56  
 30: 11 31: 11 32: 11 35: 11 36: 11 44: 87  
 48: 87

10 There are 16 hits at base# 11

BpmI ctccag 19  
 15: 12 16: 12 17: 12 18: 12 20: 12 21: 12  
 22: 12 23: 12 24: 12 25: 12 26: 12 27: 12  
 15 28: 12 30: 12 31: 12 32: 12 34: 12 35: 12  
 36: 12

There are 19 hits at base# 12

XmnI GAANNnnttc 12  
 20 37: 30 38: 30 39: 30 40: 30 41: 30 42: 30  
 43: 30 44: 30 45: 30 46: 30 47: 30 50: 30

There are 12 hits at base# 30

BsrI NCcagt 12  
 25 37: 32 38: 32 39: 32 40: 32 41: 32 42: 32  
 43: 32 44: 32 45: 32 46: 32 47: 32 50: 32

There are 12 hits at base# 32

BanII GRGCYc 11  
 30 37: 51 38: 51 39: 51 40: 51 41: 51 42: 51  
 43: 51 44: 51 45: 51 46: 51 47: 51

There are 11 hits at base# 51

Ecl136I GAGctc 11  
 35 37: 51 38: 51 39: 51 40: 51 41: 51 42: 51  
 43: 51 44: 51 45: 51 46: 51 47: 51

There are 11 hits at base# 51

SacI GAGCTc 11

There are 11 hits at base# 51

5

[illegible]

Table 206: Synthetic 3-23 FR3 of human heavy chains showing positions of possible cleavage sites

```
! Sites engineered into the synthetic gene are shown in upper case DNA
! with the RE name between vertical bars (as in | XbaI |).
! RERSs frequently found in GLGs are shown below the synthetic sequence
! with the name to the right (as in gtn ac=MaeIII(24), indicating that
! 24 of the 51 GLGs contain the site).
```

10

! Allowed DNA

15

20

25

30

35

40

45

```
|---FR3---|
89 90 (codon # in
R F synthetic 3-23)
|cgc|ttc| 6
|cgn|tty|
|agr|

ga ntc = HinfI(38)
ga gtc = PstI(18)
ga wtc = TfiI(20)

gtn ac = MaeII(24)
gts acc = Tsp45I(21)
tc acc = HphI(44)
```

-----FR3-----

```

!          91  92  93  94  95  96  97  98  99 100 101 102 103 104 105
!          T   I   S   R   D   N   S   K   N   T   L   Y   L   Q   M
!          |act|atc|TCT|AGA|gac|aac|tct|aag|aat|atc|ctc|tacc|ttg|cag|atg|  51
!allowed|acn|ath|tcn|cgn|gay|aay|tcn|aar|aay|acn|ttr|tay|ttr|car|atg|
!          |agy|agr|          |agy|          |ctn|          |ctn|
!          |          |gal|gac = BsmAI(16)          |          |ag ct = AluI(23)
!          |          |c|tcc ag = BpmI(19)          |          |g ctn agc = BlpI(21)
!          |          |          |          |          |g aan nnn ttc = XmnI(12)
!          |          |          |          |          |tg ca = HpyCH4V(21)
!          |          |XbaI|          |          |          |          |

```

-----FR3----->

[illegible]

Table 217: Human HC GLG FR1 Sequences

VH Exon - Nucleotide sequence alignment

VH1

	1-02	CAG GTG CAG CTG GTG CAG TCT GGG GCT GAG GTG AAG AAG CCT GGG GCC TCA GTG AAG GTC TCC TGC AAG GCT TCT GGA TAC ACC TTC ACC
5	1-03	cag gtC cag ctT gtg cag tct ggg gct gag gtg aag aag cct ggg gcc tca gtg aag gtT tcc tgc aag gct tct gga tac acc ttc acT
	1-08	cag gtg cag ctg gtg cag tct ggg gct gag gtg aag aag cct ggg gcc tca gtg aag gtc tcc tgc aag gct tct gga tac acc ttc acc
10	1-18	cag gtT cag ctg gtg cag tct ggA gct gag gtg aag aag cct ggg gcc tca gtg aag gtc tcc tgc aag gct tct ggT tac acc ttT acc
	1-24	cag gtC cag ctg gtA cag tct ggg gct gag gtg aag aag cct ggg gcc tca gtg aag gtc tcc tgc aag gTt tcC gga tac acc Ctc acT
	1-45	cag Atg cag ctg gtg cag tct ggg gct gag gtg aag aag Act ggg Tcc tca gtg aag gtT tcc tgc aag gct tcC gga tac acc ttc acc
15	1-46	cag gtg cag ctg gtg cag tct ggg gct gag gtg aag aag cct ggg gcc tca gtg aag gtT tcc tgc aag gcA tct gga tac acc ttc acc
	1-58	caA Atg cag ctg gtg cag tct ggg Cct gag gtg aag aag cct ggg Acc tca gtg aag gtc tcc tgc aag gct tct gga tTc acc ttT acT
20	1-69	cag gtg cag ctg gtg cag tct ggg gct gag gtg aag aag cct ggg Tcc tcG gtg aag gtc tcc tgc aag gct tct gga GGc acc ttc aGc
	1-e	cag gtg cag ctg gtg cag tct ggg gct gag gtg aag aag cct ggg Tcc tcG gtg aag gtc tcc tgc aag gct tct gga GGc acc ttc aGc
	1-f	Gag gtC cag ctg gtA cag tct ggg gct gag gtg aag aag cct ggg gcT Aca gtg aaA Atc tcc tgc aag gTt tct gga tac acc ttc acc
25	VH2	
	2-05	CAG ATC ACC TTG AAG GAG TCT GGT CCT ACG CTG GTG AAA CCC ACA CAG ACC CTC ACG CTG ACC TGC ACC TTC TCT GGG TTC TCA CTC AGC
	2-26	cag Gtc acc ttg aag gag tct ggt cct GTg ctg gtg aaa ccc aca Gag acc ctc acg ctg acc tgc acc Gtc tct ggg ttc tca ctc agc
30	2-70	cag Gtc acc ttg aag gag tct ggt cct Gcg ctg gtg aaa ccc aca cag acc ctc acA ctg acc tgc acc ttc tct ggg ttc tca ctc agc
	VH3	
	3-07	GAG GTG CAG CTG GTG GAG TCT GGG GGA GGC TTG GTC CAG CCT GGG GGG TCC CTG AGA CTC TCC TGT GCA GCC TCT GGA TTC ACC TTT AGT
35	3-09	gaA gtg cag ctg gtg gag tct ggg gga ggc ttg gtA cag cct ggC Agg tcc ctg aga ctc tcc tgt gca gcc tct gga ttc acc ttt GAt
	3-11	Cag gtg cag ctg gtg gag tct ggg gga ggc ttg gtc Aag cct ggA ggg tcc ctg aga ctc tcc tgt gca gcc tct gga ttc acc ttC agt
40	3-13	gag gtg cag ctg gtg gag tct ggg gga ggc ttg gtA cag cct ggg ggg tcc ctg aga ctc tcc tgt gca gcc tct gga ttc acc ttC agt
	3-15	gag gtg cag ctg gtg gag tct ggg gga ggc ttg gtA Aag cct ggg ggg tcc ctT aga ctc tcc tgt gca gcc tct gga ttc acT ttC agt
	3-20	gag gtg cag ctg gtg gag tct ggg gga ggT Gtg gtA cGg cct ggg ggg tcc ctg aga

ctc tcc tgt gca gcc tct gga ttc acc ttt GAt  
 5 3-21 gag gtg cag ctg gtg gag tct ggg gga ggc Ctg gtc Aag cct ggg ggg tcc ctg aga  
 ctc tcc tgt gca gcc tct gga ttc acc ttC agt  
 3-23 gag gtg cag ctg Ttg gag tct ggg gga ggc ttg gtA cag cct ggg ggg tcc ctg aga  
 ctc tcc tgt gca gcc tct gga ttc acc ttt agC  
 3-30 Cag gtg cag ctg gtg gag tct ggg gga ggc Gtg gtc cag cct ggg Agg tcc ctg aga  
 ctc tcc tgt gca gcc tct gga ttc acc ttC agt  
 3-30.3 Cag gtg cag ctg gtg gag tct ggg gga ggc Gtg gtc cag cct ggg Agg tcc ctg aga  
 ctc tcc tgt gca gcc tct gga ttc acc ttC agt  
 10 3-30.5 Cag gtg cag ctg gtg gag tct ggg gga ggc Gtg gtc cag cct ggg Agg tcc ctg aga  
 ctc tcc tgt gca gcc tct gga ttc acc ttC agt  
 3-33 Cag gtg cag ctg gtg gag tct ggg gga ggc Gtg gtc cag cct ggg Agg tcc ctg aga  
 ctc tcc tgt gca gcc tct gga ttc acc ttC agt  
 15 3-43 gaA gtg cag ctg gtg gag tct ggg gga gTc Gtg gtA cag cct ggg ggg tcc ctg aga  
 ctc tcc tgt gca gcc tct gga ttc acc ttt GAt  
 3-48 gag gtg cag ctg gtg gag tct ggg gga ggc ttg gtA cag cct ggg ggg tcc ctg aga  
 ctc tcc tgt gca gcc tct gga ttc acc ttC agt  
 3-49 gag gtg cag ctg gtg gag tct ggg gga ggc ttg gtA cag cCA ggg Cgg tcc ctg aga  
 ctc tcc tgt Aca gCT tct gga ttc acc ttt Ggt  
 20 3-53 gag gtg cag ctg gtg gag Act ggA gga ggc ttg Atc cag cct ggg ggg tcc ctg aga  
 ctc tcc tgt gca gcc tct ggG ttc acc GtC agt  
 3-64 gag gtg cag ctg gtg gag tct ggg gga ggc ttg gtc cag cct ggg ggg tcc ctg aga  
 ctc tcc tgt gca gcc tct gga ttc acc ttC agt  
 25 3-66 gag gtg cag ctg gtg gag tct ggg gga ggc ttg gtc cag cct ggg ggg tcc ctg aga  
 ctc tcc tgt gca gcc tct gga ttc acc GtC agt  
 3-72 gag gtg cag ctg gtg gag tct ggg gga ggc ttg gtc cag cct ggA ggg tcc ctg aga  
 ctc tcc tgt gca gcc tct gga ttc acc ttC agt  
 3-73 gag gtg cag ctg gtg gag tct ggg gga ggc ttg gtc cag cct ggg ggg tcc ctg aAa  
 ctc tcc tgt gca gcc tct ggG ttc acc ttC agt  
 30 3-74 gag gtg cag ctg gtg gag tCg ggg gga ggc tTA gtT cag cct ggg ggg tcc ctg aga  
 ctc tcc tgt gca gcc tct gga ttc acc ttC agt  
 3-d gag gtg cag ctg gtg gag tct Cgg gga gTc ttg gtA cag cct ggg ggg tcc ctg aga  
 ctc tcc tgt gca gcc tct gga ttc acc GtC agt  
 VH4  
 35 4-04 CAG GTG CAG CTG CAG GAG TCG GGC CCA GGA CTG GTG AAG CCT TCG GGG ACC CTG TCC  
 CTC ACC TGC GCT GTC TCT GGT GGC TCC ATC AGC  
 4-28 cag gtg cag ctg cag gag tCg ggc cca gga ctg gtg aag cct tCg gAC acc ctg tcc  
 ctc acc tgc gct gtc tct ggt TAc tcc atc agc  
 4-30.1 cag gtg cag ctg cag gag tCg ggc cca gga ctg gtg aag cct tCA CAg acc ctg tcc  
 ctc acc tgc Act gtc tct ggt ggc tcc atc agc  
 40 4-30.2 cag Ctg cag ctg cag gag tCg ggc Tca gga ctg gtg aag cct tCA CAg acc ctg tcc  
 ctc acc tgc gct gtc tct ggt ggc tcc atc agc  
 4-30.4 cag gtg cag ctg cag gag tCg ggc cca gga ctg gtg aag cct tCA CAg acc ctg tcc  
 ctc acc tgc Act gtc tct ggt ggc tcc atc agc



Table 220: RERS sites in Human HC GLG FR1s where there are at least 20 GLGs cut

BsgI GTGCAG 71 (cuts 16/14 bases to right)

	1: 4	1: 13	2: 13	3: 4	3: 13	4: 13
	6: 13	7: 4	7: 13	8: 13	9: 4	9: 13
5	10: 4	10: 13	15: 4	15: 65	16: 4	16: 65
	17: 4	17: 65	18: 4	18: 65	19: 4	19: 65
	20: 4	20: 65	21: 4	21: 65	22: 4	22: 65
	23: 4	23: 65	24: 4	24: 65	25: 4	25: 65
	26: 4	26: 65	27: 4	27: 65	28: 4	28: 65
10	29: 4	30: 4	30: 65	31: 4	31: 65	32: 4
	32: 65	33: 4	33: 65	34: 4	34: 65	35: 4
	35: 65	36: 4	36: 65	37: 4	38: 4	39: 4
	41: 4	42: 4	43: 4	45: 4	46: 4	47: 4
	48: 4	48: 13	49: 4	49: 13	51: 4	

15 There are 39 hits at base# 4

There are 21 hits at base# 65

-"	ctgcac	9			
	12: 63	13: 63	14: 63	39: 63	41: 63 42: 63
20	44: 63	45: 63	46: 63		

	BbvI GCAGC	65			
	1: 6	3: 6	6: 6	7: 6	8: 6 9: 6
	10: 6	15: 6	15: 67	16: 6	16: 67 17: 6
	17: 67	18: 6	18: 67	19: 6	19: 67 20: 6
25	20: 67	21: 6	21: 67	22: 6	22: 67 23: 6
	23: 67	24: 6	24: 67	25: 6	25: 67 26: 6
	26: 67	27: 6	27: 67	28: 6	28: 67 29: 6
	30: 6	30: 67	31: 6	31: 67	32: 6 32: 67
	33: 6	33: 67	34: 6	34: 67	35: 6 35: 67
30	36: 6	36: 67	37: 6	38: 6	39: 6 40: 6
	41: 6	42: 6	43: 6	44: 6	45: 6 46: 6
	47: 6	48: 6	49: 6	50: 12	51: 6

There are 43 hits at base# 6 Bolded sites very near sites listed below

35 There are 21 hits at base# 67

-"	gctgc	13			
	37: 9	38: 9	39: 9	40: 3	40: 9 41: 9
	42: 9	44: 3	44: 9	45: 9	46: 9 47: 9

50: 9

There are 11 hits at base# 9

BsoFI GCngc

78

5 1: 6 3: 6 6: 6 7: 6 8: 6 9: 6  
10: 6 15: 6 15: 67 16: 6 16: 67 17: 6  
17: 67 18: 6 18: 67 19: 6 19: 67 20: 6  
20: 67 21: 6 21: 67 22: 6 22: 67 23: 6  
23: 67 24: 6 24: 67 25: 6 25: 67 26: 6  
10 26: 67 27: 6 27: 67 28: 6 28: 67 29: 6  
30: 6 30: 67 31: 6 31: 67 32: 6 32: 67  
33: 6 33: 67 34: 6 34: 67 35: 6 35: 67  
36: 6 36: 67 37: 6 37: 9 38: 6 38: 9  
39: 6 39: 9 40: 3 40: 6 40: 9 41: 6  
15 41: 9 42: 6 42: 9 43: 6 44: 3 44: 6  
44: 9 45: 6 45: 9 46: 6 46: 9 47: 6  
47: 9 48: 6 49: 6 50: 9 50: 12 51: 6

There are 43 hits at base# 6 These often occur together.

There are 11 hits at base# 9

20 There are 2 hits at base# 3

There are 21 hits at base# 67

TseI GcwgC

78

25 1: 6 3: 6 6: 6 7: 6 8: 6 9: 6  
10: 6 15: 6 15: 67 16: 6 16: 67 17: 6  
17: 67 18: 6 18: 67 19: 6 19: 67 20: 6  
20: 67 21: 6 21: 67 22: 6 22: 67 23: 6  
23: 67 24: 6 24: 67 25: 6 25: 67 26: 6  
26: 67 27: 6 27: 67 28: 6 28: 67 29: 6  
30 30: 6 30: 67 31: 6 31: 67 32: 6 32: 67  
33: 6 33: 67 34: 6 34: 67 35: 6 35: 67  
36: 6 36: 67 37: 6 37: 9 38: 6 38: 9  
39: 6 39: 9 40: 3 40: 6 40: 9 41: 6  
41: 9 42: 6 42: 9 43: 6 44: 3 44: 6  
35 44: 9 45: 6 45: 9 46: 6 46: 9 47: 6  
47: 9 48: 6 49: 6 50: 9 50: 12 51: 6

There are 43 hits at base# 6 Often together.

There are 11 hits at base# 9



There are 2 hits at base# 3  
 There are 1 hits at base# 12  
 There are 21 hits at base# 67

5 MspAII CMGckg 48  
 1: 7 3: 7 4: 7 5: 7 6: 7 7: 7  
 8: 7 9: 7 10: 7 11: 7 15: 7 16: 7  
 17: 7 18: 7 19: 7 20: 7 21: 7 22: 7  
 23: 7 24: 7 25: 7 26: 7 27: 7 28: 7  
 10 29: 7 30: 7 31: 7 32: 7 33: 7 34: 7  
 35: 7 36: 7 37: 7 38: 7 39: 7 40: 1  
40: 7 41: 7 42: 7 44: 1 44: 7 45: 7  
 46: 7 47: 7 48: 7 49: 7 50: 7 51: 7  
 There are 46 hits at base# 7

15 PvuII CAGctg 48  
 1: 7 3: 7 4: 7 5: 7 6: 7 7: 7  
 8: 7 9: 7 10: 7 11: 7 15: 7 16: 7  
 17: 7 18: 7 19: 7 20: 7 21: 7 22: 7  
 20 23: 7 24: 7 25: 7 26: 7 27: 7 28: 7  
 29: 7 30: 7 31: 7 32: 7 33: 7 34: 7  
 35: 7 36: 7 37: 7 38: 7 39: 7 40: 1  
40: 7 41: 7 42: 7 44: 1 44: 7 45: 7  
 46: 7 47: 7 48: 7 49: 7 50: 7 51: 7  
 25 There are 46 hits at base# 7  
 There are 2 hits at base# 1

AluI AGct 54  
 1: 8 2: 8 3: 8 4: 8 4: 24 5: 8  
 30 6: 8 7: 8 8: 8 9: 8 10: 8 11: 8  
 15: 8 16: 8 17: 8 18: 8 19: 8 20: 8  
 21: 8 22: 8 23: 8 24: 8 25: 8 26: 8  
 27: 8 28: 8 29: 8 29: 69 30: 8 31: 8  
 32: 8 33: 8 34: 8 35: 8 36: 8 37: 8  
 35 38: 8 39: 8 40: 2 40: 8 41: 8 42: 8  
 43: 8 44: 2 44: 8 45: 8 46: 8 47: 8  
 48: 8 48: 82 49: 8 49: 82 50: 8 51: 8  
 There are 48 hits at base# 8

There are 2 hits at base# 2

DdeI Ctnag

48

5 1: 26 1: 48 2: 26 2: 48 3: 26 3: 48  
4: 26 4: 48 5: 26 5: 48 6: 26 6: 48  
7: 26 7: 48 8: 26 8: 48 9: 26 10: 26  
11: 26 12: 85 13: 85 14: 85 15: 52 16: 52  
17: 52 18: 52 19: 52 20: 52 21: 52 22: 52  
23: 52 24: 52 25: 52 26: 52 27: 52 28: 52  
10 29: 52 30: 52 31: 52 32: 52 33: 52 35: 30  
35: 52 36: 52 40: 24 49: 52 51: 26 51: 48

There are 22 hits at base# 52 52 and 48 never together.

There are 9 hits at base# 48

There are 12 hits at base# 26 26 and 24 never together.

15

HphI tcacc

42

1: 86 3: 86 6: 86 7: 86 8: 80 11: 86  
12: 5 13: 5 14: 5 15: 80 16: 80 17: 80  
18: 80 20: 80 21: 80 22: 80 23: 80 24: 80  
20 25: 80 26: 80 27: 80 28: 80 29: 80 30: 80  
31: 80 32: 80 33: 80 34: 80 35: 80 36: 80  
37: 59 38: 59 39: 59 40: 59 41: 59 42: 59  
43: 59 44: 59 45: 59 46: 59 47: 59 50: 59

There are 22 hits at base# 80 80 and 86 never together

25

There are 5 hits at base# 86

There are 12 hits at base# 59

BssKI Nccngg

50

1: 39 2: 39 3: 39 4: 39 5: 39 7: 39  
30 8: 39 9: 39 10: 39 11: 39 15: 39 16: 39  
17: 39 18: 39 19: 39 20: 39 21: 29 21: 39  
22: 39 23: 39 24: 39 25: 39 26: 39 27: 39  
28: 39 29: 39 30: 39 31: 39 32: 39 33: 39  
34: 39 35: 19 35: 39 36: 39 37: 24 38: 24  
35 39: 24 41: 24 42: 24 44: 24 45: 24 46: 24  
47: 24 48: 39 48: 40 49: 39 49: 40 50: 24  
50: 73 51: 39

There are 35 hits at base# 39 39 and 40 together twice.

There are 2 hits at base# 40

BsaJI Ccnngg

47

1: 40 2: 40 3: 40 4: 40 5: 40 7: 40  
8: 40 9: 40 9: 47 10: 40 10: 47 11: 40  
5 15: 40 18: 40 19: 40 20: 40 21: 40 22: 40  
23: 40 24: 40 25: 40 26: 40 27: 40 28: 40  
29: 40 30: 40 31: 40 32: 40 34: 40 35: 20  
35: 40 36: 40 37: 24 38: 24 39: 24 41: 24  
42: 24 44: 24 45: 24 46: 24 47: 24 48: 40  
10 48: 41 49: 40 49: 41 50: 74 51: 40

There are 32 hits at base# 40 40 and 41 together twice

There are 2 hits at base# 41

There are 9 hits at base# 24

There are 2 hits at base# 47

15

BstNI CCwgg

44

PspGI ccwgg

ScrFI(\$M.HpaII) CCwgg

1: 40 2: 40 3: 40 4: 40 5: 40 7: 40  
20 8: 40 9: 40 10: 40 11: 40 15: 40 16: 40  
17: 40 18: 40 19: 40 20: 40 21: 30 21: 40  
22: 40 23: 40 24: 40 25: 40 26: 40 27: 40  
28: 40 29: 40 30: 40 31: 40 32: 40 33: 40  
34: 40 35: 40 36: 40 37: 25 38: 25 39: 25  
25 41: 25 42: 25 44: 25 45: 25 46: 25 47: 25  
50: 25 51: 40

There are 33 hits at base# 40

ScrFI CCngg

50

30 1: 40 2: 40 3: 40 4: 40 5: 40 7: 40  
8: 40 9: 40 10: 40 11: 40 15: 40 16: 40  
17: 40 18: 40 19: 40 20: 40 21: 30 21: 40  
22: 40 23: 40 24: 40 25: 40 26: 40 27: 40  
28: 40 29: 40 30: 40 31: 40 32: 40 33: 40  
35 34: 40 35: 20 35: 40 36: 40 37: 25 38: 25  
39: 25 41: 25 42: 25 44: 25 45: 25 46: 25  
47: 25 48: 40 48: 41 49: 40 49: 41 50: 25  
50: 74 51: 40

There are 35 hits at base# 40

There are 2 hits at base# 41

EcoO109I RGnccy

34

5 1: 43 2: 43 3: 43 4: 43 5: 43 6: 43  
7: 43 8: 43 9: 43 10: 43 15: 46 16: 46  
17: 46 18: 46 19: 46 20: 46 21: 46 22: 46  
23: 46 24: 46 25: 46 26: 46 27: 46 28: 46  
30: 46 31: 46 32: 46 33: 46 34: 46 35: 46  
36: 46 37: 46 43: 79 51: 43

10 There are 22 hits at base# 46 46 and 43 never together

There are 11 hits at base# 43

NlaIV GGNncc

71

15 1: 43 2: 43 3: 43 4: 43 5: 43 6: 43  
7: 43 8: 43 9: 43 9: 79 10: 43 10: 79  
15: 46 15: 47 16: 47 17: 46 17: 47 18: 46  
18: 47 19: 46 19: 47 20: 46 20: 47 21: 46  
21: 47 22: 46 22: 47 23: 47 24: 47 25: 47  
26: 47 27: 46 27: 47 28: 46 28: 47 29: 47  
30: 46 30: 47 31: 46 31: 47 32: 46 32: 47  
20 33: 46 33: 47 34: 46 34: 47 35: 46 35: 47  
36: 46 36: 47 37: 21 37: 46 37: 47 37: 79  
38: 21 39: 21 39: 79 40: 79 41: 21 41: 79  
42: 21 42: 79 43: 79 44: 21 44: 79 45: 21  
45: 79 46: 21 46: 79 47: 21 51: 43

25 There are 23 hits at base# 47 46 & 47 often together

There are 17 hits at base# 46

There are 11 hits at base# 43

Sau96I Ggncc

70

30 1: 44 2: 3 2: 44 3: 44 4: 44 5: 3 5: 44 6: 44  
7: 44 8: 22 8: 44 9: 44 10: 44 11: 3 12: 22 13: 22  
14: 22 15: 33 15: 47 16: 47 17: 47 18: 47 19: 47 20: 47  
21: 47 22: 47 23: 33 23: 47 24: 33 24: 47 25: 33 25: 47  
26: 33 26: 47 27: 47 28: 47 29: 47 30: 47 31: 33 31: 47  
32: 33 32: 47 33: 33 33: 47 34: 33 34: 47 35: 47 36: 47  
37: 21 37: 22 37: 47 38: 21 38: 22 39: 21 39: 22 41: 21  
35 41: 22 42: 21 42: 22 43: 80 44: 21 44: 22 45: 21 45: 22  
46: 21 46: 22 47: 21 47: 22 50: 22 51: 44

There are 23 hits at base# 47 These do not occur together.

There are 11 hits at base# 44

There are 14 hits at base# 22 These do occur together.  
There are 9 hits at base# 21

BsmAI GTCTCNnnnn 22

5 1: 58 3: 58 4: 58 5: 58 8: 58 9: 58  
10: 58 13: 70 36: 18 37: 70 38: 70 39: 70  
40: 70 41: 70 42: 70 44: 70 45: 70 46: 70  
47: 70 48: 48 49: 48 50: 85

There are 11 hits at base# 70

10

-- Nnnnnngagac 27

13: 40 15: 48 16: 48 17: 48 18: 48 20: 48  
21: 48 22: 48 23: 48 24: 48 25: 48 26: 48  
27: 48 28: 48 29: 48 30: 10 30: 48 31: 48

15

32: 48 33: 48 35: 48 36: 48 43: 40 44: 40  
45: 40 46: 40 47: 40

There are 20 hits at base# 48

AvaII Ggwcc 44

20

Sau96I(\$M.HaeIII) Ggwcc 44

2: 3 5: 3 6: 44 8: 44 9: 44 10: 44  
11: 3 12: 22 13: 22 14: 22 15: 33 15: 47  
16: 47 17: 47 18: 47 19: 47 20: 47 21: 47  
22: 47 23: 33 23: 47 24: 33 24: 47 25: 33  
25: 47 26: 33 26: 47 27: 47 28: 47 29: 47  
30: 47 31: 33 31: 47 32: 33 32: 47 33: 33  
33: 47 34: 33 34: 47 35: 47 36: 47 37: 47  
43: 80 50: 22

25

There are 23 hits at base# 47 44 & 47 never together

30

There are 4 hits at base# 44

PpuMI RGgwccy 27

35

6: 43 8: 43 9: 43 10: 43 15: 46 16: 46  
17: 46 18: 46 19: 46 20: 46 21: 46 22: 46  
23: 46 24: 46 25: 46 26: 46 27: 46 28: 46  
30: 46 31: 46 32: 46 33: 46 34: 46 35: 46  
36: 46 37: 46 43: 79

There are 22 hits at base# 46 43 and 46 never occur together.

There are 4 hits at base# 43

BsmFI GGGAC

3

8: 43 37: 46 50: 77

-"- gtccc

33

5 15: 48 16: 48 17: 48 1: 0 1: 0 20: 48  
21: 48 22: 48 23: 48 24: 48 25: 48 26: 48  
27: 48 28: 48 29: 48 30: 48 31: 48 32: 48  
33: 48 34: 48 35: 48 36: 48 37: 54 38: 54  
39: 54 40: 54 41: 54 42: 54 43: 54 44: 54  
10 45: 54 46: 54 47: 54

There are 20 hits at base# 48

There are 11 hits at base# 54

HinfI Gantc

80

15 8: 77 12: 16 13: 16 14: 16 15: 16 15: 56  
15: 77 16: 16 16: 56 16: 77 17: 16 17: 56  
17: 77 18: 16 18: 56 18: 77 19: 16 19: 56  
19: 77 20: 16 20: 56 20: 77 21: 16 21: 56  
21: 77 22: 16 22: 56 22: 77 23: 16 23: 56  
20 23: 77 24: 16 24: 56 24: 77 25: 16 25: 56  
25: 77 26: 16 26: 56 26: 77 27: 16 27: 26  
27: 56 27: 77 28: 16 28: 56 28: 77 29: 16  
29: 56 29: 77 30: 56 31: 16 31: 56 31: 77  
32: 16 32: 56 32: 77 33: 16 33: 56 33: 77  
25 34: 16 35: 16 35: 56 35: 77 36: 16 36: 26  
36: 56 36: 77 37: 16 38: 16 39: 16 40: 16  
41: 16 42: 16 44: 16 45: 16 46: 16 47: 16  
48: 46 49: 46

There are 34 hits at base# 16

30

TfiI Gawtc

21

8: 77 15: 77 16: 77 17: 77 18: 77 19: 77  
20: 77 21: 77 22: 77 23: 77 24: 77 25: 77  
26: 77 27: 77 28: 77 29: 77 31: 77 32: 77  
35 33: 77 35: 77 36: 77

There are 21 hits at base# 77

MlyI GAGTC 38

12: 16	13: 16	14: 16	15: 16	16: 16	17: 16
18: 16	19: 16	20: 16	21: 16	22: 16	23: 16
24: 16	25: 16	26: 16	27: 16	27: 26	28: 16
29: 16	31: 16	32: 16	33: 16	34: 16	35: 16
36: 16	36: 26	37: 16	38: 16	39: 16	40: 16
41: 16	42: 16	44: 16	45: 16	46: 16	47: 16
48: 46	49: 46				

5

There are 34 hits at base# 16

10

--- GACTC 21

15: 56	16: 56	17: 56	18: 56	19: 56	20: 56
21: 56	22: 56	23: 56	24: 56	25: 56	26: 56
27: 56	28: 56	29: 56	30: 56	31: 56	32: 56

15

There are 21 hits at base# 56

PleI gagtc 38

12: 16	13: 16	14: 16	15: 16	16: 16	17: 16
18: 16	19: 16	20: 16	21: 16	22: 16	23: 16
24: 16	25: 16	26: 16	27: 16	27: 26	28: 16
29: 16	31: 16	32: 16	33: 16	34: 16	35: 16
36: 16	36: 26	37: 16	38: 16	39: 16	40: 16
41: 16	42: 16	44: 16	45: 16	46: 16	47: 16
48: 46	49: 46				

20

There are 34 hits at base# 16

25

--- gactc 21

15: 56	16: 56	17: 56	18: 56	19: 56	20: 56
21: 56	22: 56	23: 56	24: 56	25: 56	26: 56
27: 56	28: 56	29: 56	30: 56	31: 56	32: 56
33: 56	35: 56	36: 56			

30

There are 21 hits at base# 56

AlwNI CAGNNNctg 26

15: 68	16: 68	17: 68	18: 68	19: 68	20: 68
21: 68	22: 68	23: 68	24: 68	25: 68	26: 68
27: 68	28: 68	29: 68	30: 68	31: 68	32: 68
33: 68	34: 68	35: 68	36: 68	39: 46	40: 46
41: 46	42: 46				

35

There are 22 hits at base# 68

Table 255: Analysis of frequency of matching REadaptors in actual V genes

A: HpyCH4V in HC at bases 35-56

	Id	Ntot	Number of mismatches..... Number										Id	Probe		
			0	1	2	3	4	5	6	7	8	9			10	Cut
5	1	510	5	11	274	92	61	25	22	11	1	3	5	443	6-1	agttctcccTGCAGctgaactc
	2	192	54	42	32	24	15	2	3	10	3	1	6	167	3-11	cactgtatcTGCaaatgaacag
	3	58	19	7	17	6	5	1	0	1	0	2	0	54	3-09	ccctgtatcTGCaaatgaacag
10	4	267	42	33	9	8	8	82	43	22	8	11	1	100	5-51	ccgcctaccTGCAGtgaggcag
	5	250	111	59	41	24	7	5	1	0	0	2	0	242	3-15	cgctgtatcTGCaaatgaacag
	6	7	0	2	0	1	0	0	0	0	0	4	0	3	7-4.1	cggcatatcTGCAGatctgcag
	7	7	0	2	2	0	0	2	1	0	0	0	0	4	3-73	cggcgtatcTGCaaatgaacag
	8	26	10	4	1	3	1	2	1	3	1	0	0	19	5-a	ctgcctaccTGCAGtgaggcag
	9	21	8	2	3	1	6	1	0	0	0	0	0	20	3-49	tcgcctatcTGCaaatgaacag
15		1338	249	162	379	149	103	120	71	47	13	23	12	1052		
			249	411	790	939		1162		1280		1316				
							1042		1233		1293		1338			

Id	Probe	dotted probe									
		6-1	3-11	3-09	5-51	3-15	7-4.1	3-73	5-a	3-49	
20	agttctcccTGCAGctgaactc	agttctcccTGCAGctgaactc	cactgtatcTGCaaatgaacag	ccctgtatcTGCaaatgaacag	ccgcctaccTGCAGtgaggcag	cgctgtatcTGCaaatgaacag	cggcatatcTGCAGatctgcag	cggcgtatcTGCaaatgaacag	ctgcctaccTGCAGtgaggcag	tcgcctatcTGCaaatgaacag	
25	cactgtatcTGCaaatgaacag	cac.g.at.....aa.....ag	ccctgtatcTGCaaatgaacag	ccg.c.a.....tg..g.ag	c.c.g.at.....aa.....ag	cggcatatcTGCAGatctgcag	c.gca.at.....a.ctg.ag	cggcgtatcTGCaaatgaacag	ctgc.a.....tg..g.ag	tcgc..at.....aa.....ag	



Seqs with the expected RE site only.....1004  
 (Counts only cases with 4 or fewer mismatches)

Seqs with only an unexpected site..... 0  
 Seqs with both expected and unexpected..... 48  
 (Counts only cases with 4 or fewer mismatches)

Seqs with no sites..... 0

5

B: BpI in HC

	Id	Ntot	0	1	2	3	4	5	6	7	8	Ncut	Name
10	1	133	73	16	11	13	6	9	1	4	0	119	1-58 acatggagCTGAGCagcctgag
	2	14	11	1	0	0	0	0	1	0	1	12	1-02 acatggagctgagcaggctgag
	3	34	17	8	2	6	1	0	0	0	0	0	1-18 acatggagctgaggagcctgag
	4	120	50	32	16	10	9	1	1	1	0	2	5-51 acctgcagtgaggcagcctgaa
	5	55	13	11	10	17	3	1	0	0	0	0	3-15 atctgcaaatgaacagcctgaa
15	6	340	186	88	41	15	6	3	0	1	0	0	3303 atctgcaaatgaacagcctgag
	7	82	25	16	25	12	1	3	0	0	0	0	3-20 atctgcaaatgaacagctctgag
	8	3	0	2	0	1	0	0	0	0	0	0	74.1 atctgcagatctgcagcctaaa
	9	23	18	2	2	1	0	0	0	0	0	0	3-66 atcttcaaatgaacagcctgag
	10	2	1	0	1	0	0	0	0	0	0	0	3-64 atcttcaaatgggcagcctgag
20	11	486	249	78	81	38	21	10	4	4	1	467	4301 ccctgaagctgagctctgtgac
	12	16	6	3	1	0	1	1	3	1	0	1	6-1 ccctgcagctgaactctgtgac
	13	28	15	8	2	2	1	0	0	0	0	0	2-70 tccttacaatgaccaacatgga
	14	2	0	2	0	0	0	0	0	0	0	0	2-26 tccttaccatgaccaacatgga
													601

Name	Full sequence	Dot mode
1-58	acatggaGCTGAGCagcctgag	acatggaGCTGAGCagcctgag
1-02	acatgga <b>gctgag</b> caggctgag	.....g.....
1-18	acatggagctgaggagcctgag	.....g.....
5-51	acctgcagtggagcagcctgaa	..C..C..tg.....a
3-15	atctgcaaatgaacagcctgaa	.tc..C.aa...a.....a
3-30.3	atctgcaaatgaacagcctgag	.tc..C.aa...a.....
3-20	atctgcaaatgaacagctctgag	.tc..C.aa...a...t....
7-4.1	atctgcagatctgcagcctaaa	.tc..C..a.ct.....a.a
3-66	atctcaaatgaacagcctgag	.tc.tc.aa...a.....
3-64	atctcaaatgggcagcctgag	.tc.tc.aa..g.....
4-30.1	ccctgaagctgagctctgtgac	C.C..a.....tctg...C
6-1	ccctgcagctgaactctgtgac	C.C..C.....a.tctg...C
2-70	tcctacaatgaccaacatgga	t.c.tacaa...C..a.a..ga
2-26	tccttaccatgaccaacatgga	t.c.tacca...C..a.a..ga

Seqs with the expected RE site only..... 597 (counting sequences with 4 or fewer mismatches)

Seqs with only an unexpected site..... 2

Seqs with both expected and unexpected.... 2

Seqs with no sites..... 686

C: HpyCH4III, Bst4CI, or TaaI in HC

In scoring whether the RE site of interest is present, only ONs that have 4 or fewer mismatches are counted.

Number of sequences..... 1617

Id	Ntot	0	1	2	3	4	5	6	7	8	Neut	acnqt	acnqt
1	244	78	92	43	18	10	1	2	0	0	241	102#1,1	ccgtgtattACTGTgcgagaga
2	457	69	150	115	66	34	11	8	3	1	434	103#2,3	ctgtgtattactgtgcgagaga
3	173	52	45	36	22	14	3	0	0	1	169	108#3	ccgtgtattactgtgcgagagg
5	4	16	0	3	2	2	1	6	0	1	8	124#5,1	ccgtgtattactgtgcaacaga
	4	0	0	1	0	1	1	0	1	0	2	145#6	ccatgtattactgtgcaagata
6	15	1	0	1	0	6	4	1	1	1	8	158#8	ccgtgtattactgtgaggcaga
7	23	4	8	5	2	2	1	1	0	0	21	205#12	ccacatattactgtgcacacag
8	9	1	1	1	0	3	2	1	0	0	6	226#13	ccacatattactgtgcacggat
10	7	1	3	1	1	0	0	1	0	0	6	270#14	ccacgtattactgtgcacggat
10	23	7	3	5	5	2	1	0	0	0	22	309#16,	cctgtgtattactgtgcaaaaga
11	35	5	10	7	6	3	3	0	1	0	31	313#18,	ctgtgtattactgtgcaagaga
12	18	2	3	2	2	6	1	0	2	0	15	315#19	ccgtgtattactgtgcacacaga
13	3	1	2	0	0	0	0	0	0	0	3	320#20	ccttgtatcactgtgcgagaga
15	117	29	23	28	22	8	4	2	1	0	110	323#22	ccgtgtattactgtgcgaaaga
15	75	21	25	13	9	1	4	2	0	0	69	330#23,	ctgtgtattactgtgcgaaaga
16	14	2	2	2	3	0	3	1	1	0	9	349#29	ccgtgtattactgtgactagaga
17	2	0	0	1	0	0	1	0	0	0	1	372#33	ccgtgtattactgtgctagaga
18	1	0	0	1	0	0	0	0	0	0	1	373#34	ccgtgtattactgtgactagaca
19	2	0	0	0	0	0	0	0	0	2	0	3d#36	ctgtgtattactgtgaagaaaga
20	34	4	9	9	4	5	3	0	0	0	31	428#38	ccgtgtattactgtgcgagaaa
21	17	5	4	2	2	3	1	0	0	0	16	4302#40	ccgtgtattactgtgcagagaga
22	75	15	17	24	7	10	1	1	0	0	73	439#44	ctgtgtattactgtgcgagagaca
23	40	14	15	4	5	1	0	1	0	0	39	551#48	ccatgtattactgtgcgagagaca
25	24	213	26	56	60	42	20	7	2	0	204	5a#49	ccatgtattactgtgcgagagaaa
Group		337	471	363	218	130	58	23	11	6			
Cumulative		337	808	1171	1389	1519	1577	1600	1611	1617			

Seqs with the expected RE site only.....1511

Seqs with only an unexpected site..... 0

# Table 255 D

Seqs with both expected and unexpected.... 8  
Seqs with no sites..... 0

Analysis repeated using only 8 best REadaptors

5	Id	Ntot	0	1	2	3	4	5	6	7	8+		
	1	301	78	101	54	32	16	9	10	1	0	281	102#1 ccgtgtattactgtgcgagaga
	2	493	69	155	125	73	37	14	11	3	6	459	103#2 ctgtgtattactgtgcgagaga
	3	189	52	45	38	23	18	5	4	1	3	176	108#3 ccgtgtattactgtgcgagagg
	4	127	29	23	28	24	10	6	5	2	0	114	323#22 ccgtatattactgtgcgaaaga
10	5	78	21	25	14	11	1	4	2	0	0	72	330#23 ctgtgtattactgtgcgaaaga
	6	79	15	17	25	8	11	1	2	0	0	76	439#44 ctgtgtattactgtgcgagaca
	7	43	14	15	5	5	3	0	1	0	0	42	551#48 ccatgtattactgtgcgagaca
	8	307	26	63	72	51	38	24	14	13	6	250	5a#49 ccatgtattactgtgcgaga
	1	102#1											ccgtgtattactgtgcgagaga
15	2	103#2											ctgtgtattactgtgcgagaga .t.....
	3	108#3											ccgtgtattactgtgcgagagg .....g
	4	323#22											ccgtatattactgtgcgaaaga ....a.....a...
	5	330#23											ctgtgtattactgtgcgaaaga .t.....a...
	6	439#44											ctgtgtattactgtgcgagaca .t.....c.
20	7	551#48											ccatgtattactgtgcgagaca ..a.....c.
	8	5a#49											ccatgtattactgtgcgagaAA ..a.....AA

Seqs with the expected RE site only.....1463 / 1617

Seqs with only an unexpected site..... 0

25 Seqs with both expected and unexpected.... 7

Seqs with no sites..... 0

Table 300: Kappa FR1 GLGs

	!	1	2	3	4	5	6	7	8	9	10	11	12	
		GAC	ATC	CAG	ATG	ACC	CAG	TCT	CCA	TCC	TCC	CTG	TCT	
	!	13	14	15	16	17	18	19	20	21	22	23		
5		GCA	TCT	GTA	GGA	GAC	AGA	GTC	ACC	ATC	ACT	TGC	!	O12
		GAC	ATC	CAG	ATG	ACC	CAG	TCT	CCA	TCC	TCC	CTG	TCT	
		GCA	TCT	GTA	GGA	GAC	AGA	GTC	ACC	ATC	ACT	TGC	!	O2
		GAC	ATC	CAG	ATG	ACC	CAG	TCT	CCA	TCC	TCC	CTG	TCT	
		GCA	TCT	GTA	GGA	GAC	AGA	GTC	ACC	ATC	ACT	TGC	!	O18
10		GAC	ATC	CAG	ATG	ACC	CAG	TCT	CCA	TCC	TCC	CTG	TCT	
		GCA	TCT	GTA	GGA	GAC	AGA	GTC	ACC	ATC	ACT	TGC	!	O8
		GAC	ATC	CAG	ATG	ACC	CAG	TCT	CCA	TCC	TCC	CTG	TCT	
		GCA	TCT	GTA	GGA	GAC	AGA	GTC	ACC	ATC	ACT	TGC	!	A20
		GAC	ATC	CAG	ATG	ACC	CAG	TCT	CCA	TCC	TCC	CTG	TCT	
15		GCA	TCT	GTA	GGA	GAC	AGA	GTC	ACC	ATC	ACT	TGC	!	A30
		AAC	ATC	CAG	ATG	ACC	CAG	TCT	CCA	TCT	GCC	ATG	TCT	
		GCA	TCT	GTA	GGA	GAC	AGA	GTC	ACC	ATC	ACT	TGT	!	L14
		GAC	ATC	CAG	ATG	ACC	CAG	TCT	CCA	TCC	TCA	CTG	TCT	
		GCA	TCT	GTA	GGA	GAC	AGA	GTC	ACC	ATC	ACT	TGT	!	L1
20		GAC	ATC	CAG	ATG	ACC	CAG	TCT	CCA	TCC	TCA	CTG	TCT	
		GCA	TCT	GTA	GGA	GAC	AGA	GTC	ACC	ATC	ACT	TGT	!	L15
		GCC	ATC	CAG	TTG	ACC	CAG	TCT	CCA	TCC	TCC	CTG	TCT	
		GCA	TCT	GTA	GGA	GAC	AGA	GTC	ACC	ATC	ACT	TGC	!	L4
		GCC	ATC	CAG	TTG	ACC	CAG	TCT	CCA	TCC	TCC	CTG	TCT	
25		GCA	TCT	GTA	GGA	GAC	AGA	GTC	ACC	ATC	ACT	TGC	!	L18
		GAC	ATC	CAG	ATG	ACC	CAG	TCT	CCA	TCT	TCC	GTG	TCT	
		GCA	TCT	GTA	GGA	GAC	AGA	GTC	ACC	ATC	ACT	TGT	!	L5
		GAC	ATC	CAG	ATG	ACC	CAG	TCT	CCA	TCT	TCT	GTG	TCT	
		GCA	TCT	GTA	GGA	GAC	AGA	GTC	ACC	ATC	ACT	TGT	!	L19
30		GAC	ATC	CAG	TTG	ACC	CAG	TCT	CCA	TCC	TTC	CTG	TCT	
		GCA	TCT	GTA	GGA	GAC	AGA	GTC	ACC	ATC	ACT	TGC	!	L8
		GCC	ATC	CGG	ATG	ACC	CAG	TCT	CCA	TTC	TCC	CTG	TCT	
		GCA	TCT	GTA	GGA	GAC	AGA	GTC	ACC	ATC	ACT	TGC	!	L23
		GCC	ATC	CGG	ATG	ACC	CAG	TCT	CCA	TCC	TCA	TTC	TCT	
35		GCA	TCT	ACA	GGA	GAC	AGA	GTC	ACC	ATC	ACT	TGT	!	L9
		GTC	ATC	TGG	ATG	ACC	CAG	TCT	CCA	TCC	TTA	CTC	TCT	

	GCA TCT ACA GGA GAC AGA GTC ACC ATC AGT TGT !	L24
	GCC ATC CAG ATG ACC CAG TCT CCA TCC TCC CTG TCT	
	GCA TCT GTA GGA GAC AGA GTC ACC ATC ACT TGC !	L11
	GAC ATC CAG ATG ACC CAG TCT CCT TCC ACC CTG TCT	
5	GCA TCT GTA GGA GAC AGA GTC ACC ATC ACT TGC !	L12
	GAT ATT GTG ATG ACC CAG ACT CCA CTC TCC CTG CCC	
	GTC ACC CCT GGA GAG CCG GCC TCC ATC TCC TGC !	O11
	GAT ATT GTG ATG ACC CAG ACT CCA CTC TCC CTG CCC	
	GTC ACC CCT GGA GAG CCG GCC TCC ATC TCC TGC !	O1
10	GAT GTT GTG ATG ACT CAG TCT CCA CTC TCC CTG CCC	
	GTC ACC CTT GGA CAG CCG GCC TCC ATC TCC TGC !	A17
	GAT GTT GTG ATG ACT CAG TCT CCA CTC TCC CTG CCC	
	GTC ACC CTT GGA CAG CCG GCC TCC ATC TCC TGC !	A1
	GAT ATT GTG ATG ACC CAG ACT CCA CTC TCT CTG TCC	
15	GTC ACC CCT GGA CAG CCG GCC TCC ATC TCC TGC !	A18
	GAT ATT GTG ATG ACC CAG ACT CCA CTC TCT CTG TCC	
	GTC ACC CCT GGA CAG CCG GCC TCC ATC TCC TGC !	A2
	GAT ATT GTG ATG ACT CAG TCT CCA CTC TCC CTG CCC	
	GTC ACC CCT GGA GAG CCG GCC TCC ATC TCC TGC !	A19
20	GAT ATT GTG ATG ACT CAG TCT CCA CTC TCC CTG CCC	
	GTC ACC CCT GGA GAG CCG GCC TCC ATC TCC TGC !	A3
	GAT ATT GTG ATG ACC CAG ACT CCA CTC TCC TCA CCT	
	GTC ACC CTT GGA CAG CCG GCC TCC ATC TCC TGC !	A23
	GAA ATT GTG TTG ACG CAG TCT CCA GGC ACC CTG TCT	
25	TTG TCT CCA GGG GAA AGA GCC ACC CTC TCC TGC !	A27
	GAA ATT GTG TTG ACG CAG TCT CCA GCC ACC CTG TCT	
	TTG TCT CCA GGG GAA AGA GCC ACC CTC TCC TGC !	A11
	GAA ATA GTG ATG ACG CAG TCT CCA GCC ACC CTG TCT	
	GTG TCT CCA GGG GAA AGA GCC ACC CTC TCC TGC !	L2
30	GAA ATA GTG ATG ACG CAG TCT CCA GCC ACC CTG TCT	
	GTG TCT CCA GGG GAA AGA GCC ACC CTC TCC TGC !	L16
	GAA ATT GTG TTG ACA CAG TCT CCA GCC ACC CTG TCT	
	TTG TCT CCA GGG GAA AGA GCC ACC CTC TCC TGC !	L6
	GAA ATT GTG TTG ACA CAG TCT CCA GCC ACC CTG TCT	
35	TTG TCT CCA GGG GAA AGA GCC ACC CTC TCC TGC !	L20
	GAA ATT GTA ATG ACA CAG TCT CCA GCC ACC CTG TCT	



Table 302 RERS sites found in Human Kappa FR1 GLGs

	MslI	FokI	PflFI	BsrI	BsmAI	MnlI	HpyCH 4V
VKI		--> <-- -->					
O12 1-69	3	3	23	15	18	47	36
O2 101-169	103	103	123	115	118	147	136
O18 201-269	203	203	223	215	218	247	236
O8 301-369	303	303	323	315	318	347	336
A20 401-469	403	403	423	415	418	447	436
A30 501-569	503	503	523	515	518	547	536
L14 601-669	603	603		615	618	647	636
L1 701-769	703	703	723	715	718	747	736
L15 801-869	803	803	823	815	818	847	836
L4 901-969	-	903	923	906	915	947	936
L18 1001-1069	-	1003		1006	1015	1047	1036
L5 1101-1169	1103	-		1115	1118	1147	1136
L19 1201-1269	1203	1203		1215	1218	1247	1236
L8 1301-1369	-	1303	1323	1306	1315	1347	1336
L23 1401-1469	1403	1403	1408		1415	1447	1436
L9 1501-1569	1503	1503	1508	1515	1518	1547	1536
L24 1601-1669	1603		1608	1615	1618	1647	1636
L11 1701-1769	1703	1703	1723		1715	1747	1736
L12 1801-1869	1803	1803		1815	1818	1847	1836

5

10

15

20



	MslI	FokI	PflFI	BsrI	BsmAI	MnlI	HpyCH 4V
VKII		--> <-- -->					
O11 1901-1969	-	-	-	-	-	1956	-
O1 2001-2069	-	-	-	-	-	2056	-
A17 2101-2169	-	-	2112	-	2118	2156	-
A1 2201-2269	-	-	2212	-	2218	2256	-
A18 2301-2369	-	-	-	-	-	2356	-
A2 2401-2469	-	-	-	-	-	2456	-
A19 2501-2569	-	-	2512	-	2518	2556	-
A3 2601-2669	-	-	2612	-	2618	2656	-
A23 2701-2769	-	-	-	-	-	2729 2756	-
VKIII							
A27 2801-2869	-	-	2812	-	2818 2839	2860	-
A11 2901-2969	-	-	2912	-	2918 2939	2960	-
I2 3001-3069	-	-	3012	-	3018 3039	3060	-
L16 3101-3169	-	-	3112	-	3118 3139	3160	-
L6 3201-3269	-	-	3212	-	3218 3239	3260	-

5

10

15

	MslI	FokI	PflFI	BsrI	BsmAI	MnlI	HpyCH 4V
L20 3301-3369	-	--> <-- -->	3312	-	3318 3339	3360	-
L25 3401-3469	-	-	3412	-	3418 3439	3460	-
VKIV							
B3 3501-3569	3503	-	3512	3515	3518 3539	3551<	-
VKV							
B2 3601-3669	-	-	3649	-	3618 3647		-
VKVI							
A26 3701-3769	-	-	3712	-	3718		-
A10 3801-3869	-	-	3812	-	3818		-
A14 3901-3969	-	-	3912	-	3918	3930>	-

5

10

Table 302 RERS sites found in Human Kappa FR1 GLGS, continued

	SfaNI	SfcI	HinfI	MlyI	MaeIII	HphI	HpaII
				--> --> <--	Tsp45I same sites	xx38 xx56 xx62	MspI xx06 xx52
VKI							
O12 1-69	37	41	53	53	55	56	-
O2 101-169	137	141	153	153	155	156	-
O18 201-269	237	241	253	253	255	256	-

15

	SfaNI	SfiCI	HinfI	MlyI --> --> <--	MaeIII Tsp45I same sites	HphI xx38 xx56 xx62	HpaII MspI xx06 xx52
O8 301-369	337	341	353	353	355	356	-
A20 401-469	437	441	453	453	455	456	-
A30 501-569	537	541	553	553	555	556	-
L14 601-669	637	641	653	653	655	656	-
L1 701-769	737	741	753	753	755	756	-
L15 801-869	837	841	853	853	855	856	-
L4 901-969	937	941	953	953	955	956	-
L18 1001-1069	1037	1041	1053	1053	1055	1056	-
L5 1101-1169	1137	1141	1153	1153	1155	1156	-
L19 1201-1269	1237	1241	1253	1253	1255	1256	-
L8 1301-1369	1337	1341	1353	1353	1355	1356	-
L23 1401-1469	1437	1441	1453	1453	1455	1456	1406
L9 1501-1569	1537	1541	1553	1553	1555	1556	1506
L24 1601-1669	1637	1641	1653	1653	1655	1656	
L11 1701-1769	1737	1741	1753	1753	1755	1756	
L12 1801-1869	1837	1841	1853	1853	1855	1856	
VK14							
O11 1901-1969	-	-	1918	1918	1937	1938	1952
O1 2001-2069	-	-	2018	2018	2037	2038	2052
A17 2101-2169	-	-	2112	2112	2137	2138	2152
A1 2201-2269	-	-	2212	2212	2237	2238	2252

5

10

15

20

	SfaNI	SfiCI	HinfI	MlyI --> --> <--	MaeIII Tsp45I same sites	HphI xx38 xx56 xx62	HpaII MspI xx06 xx52
A18 2301-2369	-	-	2318	2318	2337	2338	2352
A2 2401-2469	-	-	2418	2418	2437	2438	2452
A19 2501-2569	-	-	2512	2512	2537	2538	2552
A3 2601-2669	-	-	2612	2612	2637	2638	2652
A23 2701-2769	-	-	2718	2718	2737	2731* 2738*	-
VKVII	-	-	-	-	-	-	-
A27 2801-2869	-	-	-	-	-	-	-
A11 2901-2969	-	-	-	-	-	-	-
L2 3001-3069	-	-	-	-	-	-	-
L16 3101-3169	-	-	-	-	-	-	-
L6 3201-3269	-	-	-	-	-	-	-
L20 3301-3369	-	-	-	-	-	-	-
L25 3401-3469	-	-	-	-	-	-	-
VKIV	-	-	-	-	-	-	-
B3 3501-3569	-	-	3525	3525	-	-	-
VKV	-	-	-	-	-	-	-
B2 3601-3669	-	-	3639	3639	-	-	-
VKVI	-	-	-	-	-	-	-
A26 3701-3769	-	-	3712	3739	3737 3755	3756 3762	-
A10 3801-3869	-	-	3812	3839	3837 3855	3856 3862	-
A14 3901-3969	-	-	3939	3939	3937 3955	3956 3962	-

5

10

15

20

1944-1945

Table 302 RERS sites found in Human Kappa FR1, continued

	BsaJI xx29 xx42 xx43	BssKI (NstNI) xx22 xx30 xx43	BpmI xx20 xx41 xx44 --> --> <--	BsrFI Cac8I NaeI NgoMI V	HaeII I	Tsp509I
VKI						
O12 1-69	-	-	-	-	-	-
O2 101-169	-	-	-	-	-	-
O18 201-269	-	-	-	-	-	-
O8 301-369	-	-	-	-	-	-
A20 401-469	-	-	-	-	-	-
A30 501-569	-	-	-	-	-	-
L14 601-669	-	-	-	-	-	-
L1 701-769	-	-	-	-	-	-
L15 801-869	-	-	-	-	-	-
L4 901-969	-	-	-	-	-	-
L18 1001-1069	-	-	-	-	-	-
L5 1101-1169	-	-	-	-	-	-
L19 1201-1269	-	-	-	-	-	-
L8 1301-1369	-	-	-	-	-	-
L23 1401-1469	-	-	-	-	-	-
L9 1501-1569	-	-	-	-	-	-
L24 1601-1669	-	-	-	-	-	-

5

10

15

20



1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 20. 21. 22. 23. 24. 25. 26. 27. 28. 29. 30. 31. 32. 33. 34. 35. 36. 37. 38. 39. 40. 41. 42. 43. 44. 45. 46. 47. 48. 49. 50. 51. 52. 53. 54. 55. 56. 57. 58. 59. 60. 61. 62. 63. 64. 65. 66. 67. 68. 69. 70. 71. 72. 73. 74. 75. 76. 77. 78. 79. 80. 81. 82. 83. 84. 85. 86. 87. 88. 89. 90. 91. 92. 93. 94. 95. 96. 97. 98. 99. 100.

|               | BsaJI<br>xx29 xx42 xx43 | BssKI (NstNI)<br>xx22 xx30 xx43 | BpmI<br>xx20 xx41 xx44<br>--> --> <-- | BsrFI<br>Cac8I<br>NaeI<br>NgOMI<br>V | HaeII<br>I | Tsp509I |
|---------------|-------------------------|---------------------------------|---------------------------------------|--------------------------------------|------------|---------|
| L25 3401-3469 | 3443                    | 3443                            | 3420 3441                             | -                                    | -          | 3403    |
| VKIV          |                         |                                 |                                       |                                      |            |         |
| B3 3501-3569  | 3529                    | 3530                            | 3520                                  | -                                    | 3554       |         |
| VKV           |                         |                                 |                                       |                                      |            |         |
| B2 3601-3669  |                         | 3643                            | 3620 3641                             | -                                    | -          |         |
| VKVI          |                         |                                 |                                       |                                      |            |         |
| A26 3701-3769 |                         | -                               | 3720                                  | -                                    | -          | 3703    |
| A10 3801-3869 |                         | -                               | 3820                                  | -                                    | -          | 3803    |
| A14 3901-3969 | 3943                    | 3943                            | 3920 3941                             | -                                    | -          | -       |



Table 400 Lambda FR1 GLG sequences

! VL1

5 CAG TCT GTG CTG ACT CAG CCA CCC TCG GTG TCT GAA  
GCC CCC AGG CAG AGG GTC ACC ATC TCC TGT ! 1a  
cag tct gtg ctg acG cag ccG ccc tcA gtg tct gGG  
gcc ccA Ggg cag agg gtc acc atc tcc tgC ! 1e  
cag tct gtg ctg act cag cca ccc tcA gCg tct gGG  
Acc ccc Ggg cag agg gtc acc atc tcT tgt ! 1c  
cag tct gtg ctg act cag cca ccc tcA gCg tct gGG  
10 Acc ccc Ggg cag agg gtc acc atc tcT tgt ! 1g  
cag tct gtg Ttg acG cag ccG ccc tcA gtg tct gCG  
gcc ccA GgA cag aAg gtc acc atc tcc tgC ! 1b

! VL2

15 CAG TCT GCC CTG ACT CAG CCT CCC TCC GCG TCC GGG  
TCT CCT GGA CAG TCA GTC ACC ATC TCC TGC ! 2c  
cag tct gcc ctg act cag cct cGc tcA gTg tcc ggg  
tct cct gga cag tca gtc acc atc tcc tgc ! 2e  
cag tct gcc ctg act cag cct Gcc tcc gTg tcT ggg  
tct cct gga cag tcG Atc acc atc tcc tgc ! 2a2  
20 cag tct gcc ctg act cag cct ccc tcc gTg tcc ggg  
tct cct gga cag tca gtc acc atc tcc tgc ! 2d  
cag tct gcc ctg act cag cct Gcc tcc gTg tcT ggg  
tct cct gga cag tcG Atc acc atc tcc tgc ! 2b2

! VL3

25 TCC TAT GAG CTG ACT CAG CCA CCC TCA GTG TCC GTG  
TCC CCA GGA CAG ACA GCC AGC ATC ACC TGC ! 3r  
tcc tat gag ctg act cag cca cTc tca gtg tcA gtg  
Gcc cTG gga cag acG gcc agG atT acc tgT ! 3j  
tcc tat gag ctg acA cag cca ccc tcG gtg tcA gtg  
30 tcc cca gga caA acG gcc agG atc acc tgc ! 3p  
tcc tat gag ctg acA cag cca ccc tcG gtg tcA gtg  
tcc cTa gga cag aTG gcc agG atc acc tgc ! 3a  
tcT tCt gag ctg act cag GAC ccT GcT gtg tcT gtg  
Gcc TTG gga cag aca gTc agG atc acA tgc ! 3l

5 tcc tat gTg ctg act cag cca ccc tca gtg tcA gtg  
 gcc cca gga Aag acG gcc agG atT acc tgT ! 3h  
 tcc tat gag ctg acA cag cTa ccc tcG gtg tcA gtg  
 tcc cca gga cag aca gcc agG atc acc tgc ! 3e  
 tcc tat gag ctg aTG cag cca ccc tcG gtg tcA gtg  
 tcc cca gga cag acG gcc agG atc acc tgc ! 3m  
 tcc tat gag ctg acA cag cca Tcc tca gtg tcA gtg  
 tcT ccG gga cag aca gcc agG atc acc tgc ! V2-19

! VL4

10 CTG CCT GTG CTG ACT CAG CCC CCG TCT GCA TCT GCC  
 TTG CTG GGA GCC TCG ATC AAG CTC ACC TGC ! 4c  
 cAg cct gtg ctg act caA TcA TcC tct gcC tct gcT  
 tCC ctg gga Tcc tcg Gtc aag ctc acc tgc ! 4a  
 cAg cTt gtg ctg act caA TcG ccC tct gcC tct gcc  
 tCC ctg gga gcc tcg Gtc aag ctc acc tgc ! 4b

15

! VL5

CAG CCT GTG CTG ACT CAG CCA CCT TCC TCC TCC GCA  
 TCT CCT GGA GAA TCC GCC AGA CTC ACC TGC ! 5e  
 cag Gct gtg ctg act cag ccG Gct tcc CTc tcT gca  
 tct cct gga gCa tca gcc agT ctc acc tgc ! 5c  
 cag cct gtg ctg act cag cca Tct tcc CAT tcT gca  
 tct Tct gga gCa tca gTc aga ctc acc tgc ! 5b

20

! VL6

AAT TTT ATG CTG ACT CAG CCC CAC TCT GTG TCG GAG  
 TCT CCG GGG AAG ACG GTA ACC ATC TCC TGC ! 6a

25

! VL7

CAG ACT GTG GTG ACT CAG GAG CCC TCA CTG ACT GTG  
 TCC CCA GGA GGG ACA GTC ACT CTC ACC TGT ! 7a  
 cag Gct gtg gtg act cag gag ccc tca ctg act gtg  
 tcc cca gga ggg aca gtc act ctc acc tgt ! 7b

30

! VL8

CAG ACT GTG GTG ACC CAG GAG CCA TCG TTC TCA GTG  
 TCC CCT GGA GGG ACA GTC ACA CTC ACT TGT ! 8a

! VL9

CAG CCT GTG CTG ACT CAG CCA CCT TCT GCA TCA GCC  
TCC CTG GGA GCC TCG GTC ACA CTC ACC TGC ! 9a

! VL10

5

CAG GCA GGG CTG ACT CAG CCA CCC TCG GTG TCC AAG  
GGC TTG AGA CAG ACC GCC ACA CTC ACC TGC ! 10a

VL9  
VL10  
VL11  
VL12  
VL13  
VL14  
VL15  
VL16  
VL17  
VL18  
VL19  
VL20  
VL21  
VL22  
VL23  
VL24  
VL25  
VL26  
VL27  
VL28  
VL29  
VL30  
VL31  
VL32  
VL33  
VL34  
VL35  
VL36  
VL37  
VL38  
VL39  
VL40  
VL41  
VL42  
VL43  
VL44  
VL45  
VL46  
VL47  
VL48  
VL49  
VL50  
VL51  
VL52  
VL53  
VL54  
VL55  
VL56  
VL57  
VL58  
VL59  
VL60  
VL61  
VL62  
VL63  
VL64  
VL65  
VL66  
VL67  
VL68  
VL69  
VL70  
VL71  
VL72  
VL73  
VL74  
VL75  
VL76  
VL77  
VL78  
VL79  
VL80  
VL81  
VL82  
VL83  
VL84  
VL85  
VL86  
VL87  
VL88  
VL89  
VL90  
VL91  
VL92  
VL93  
VL94  
VL95  
VL96  
VL97  
VL98  
VL99  
VL100

Table 405 RERSs found in human lambda FR1 GLGs

! There are 31 lambda GLGs

MlyI NnnnnnGACTC

25

1: 6 3: 6 4: 6 6: 6 7: 6 8: 6  
5 9: 6 10: 6 11: 6 12: 6 15: 6 16: 6  
20: 6 21: 6 22: 6 23: 6 23: 50 24: 6  
25: 6 25: 50 26: 6 27: 6 28: 6 30: 6  
31: 6

There are 23 hits at base# 6

10

--- GAGTCNNNNNn

1

26: 34

MwoI GCNNNNNnngc

20

15 1: 9 2: 9 3: 9 4: 9 11: 9 11: 56  
12: 9 13: 9 14: 9 16: 9 17: 9 18: 9  
19: 9 20: 9 23: 9 24: 9 25: 9 26: 9  
30: 9 31: 9

There are 19 hits at base# 9

20 HinfI Gantc

27

1: 12 3: 12 4: 12 6: 12 7: 12 8: 12  
9: 12 10: 12 11: 12 12: 12 15: 12 16: 12  
20: 12 21: 12 22: 12 23: 12 23: 46 23: 56  
24: 12 25: 12 25: 56 26: 12 26: 34 27: 12  
25 28: 12 30: 12 31: 12

There are 23 hits at base# 12

PleI gactc

25

1: 12 3: 12 4: 12 6: 12 7: 12 8: 12  
9: 12 10: 12 11: 12 12: 12 15: 12 16: 12  
30 20: 12 21: 12 22: 12 23: 12 23: 56 24: 12  
25: 12 25: 56 26: 12 27: 12 28: 12 30: 12  
31: 12

There are 23 hits at base# 12

35 --- gagtc

1

26: 34

DdeI Ctnag

32

1: 14 2: 24 3: 14 3: 24 4: 14 4: 24  
5: 24 6: 14 7: 14 7: 24 8: 14 9: 14  
5 10: 14 11: 14 11: 24 12: 14 12: 24 15: 5  
15: 14 16: 14 16: 24 19: 24 20: 14 23: 14  
24: 14 25: 14 26: 14 27: 14 28: 14 29: 30  
30: 14 31: 14

There are 21 hits at base# 14

10

BsaJI Ccnngg

38

1: 23 1: 40 2: 39 2: 40 3: 39 3: 40  
4: 39 4: 40 5: 39 11: 39 12: 38 12: 39  
13: 23 13: 39 14: 23 14: 39 15: 38 16: 39  
15 17: 23 17: 39 18: 23 18: 39 21: 38 21: 39  
21: 47 22: 38 22: 39 22: 47 26: 40 27: 39  
28: 39 29: 14 29: 39 30: 38 30: 39 30: 47  
31: 23 31: 32

There are 17 hits at base# 39

20 There are 5 hits at base# 38

There are 5 hits at base# 40 Makes cleavage ragged.

MnlI cctc

35

1: 23 2: 23 3: 23 4: 23 5: 23 6: 19  
6: 23 7: 19 8: 23 9: 19 9: 23 10: 23  
25 11: 23 13: 23 14: 23 16: 23 17: 23 18: 23  
19: 23 20: 47 21: 23 21: 29 21: 47 22: 23  
22: 29 22: 35 22: 47 23: 26 23: 29 24: 27  
27: 23 28: 23 30: 35 30: 47 31: 23

There are 21 hits at base# 23

30 There are 3 hits at base# 19

There are 3 hits at base# 29

There are 1 hits at base# 26

There are 1 hits at base# 27 These could make cleavage ragged.

-"- gagg

7

35 1: 48 2: 48 3: 48 4: 48 27: 44 28: 44

29: 44

BssKI Nccngg

39

1: 40 2: 39 3: 39 3: 40 4: 39 4: 40  
5 5: 39 6: 31 6: 39 7: 31 7: 39 8: 39  
9: 31 9: 39 10: 39 11: 39 12: 38 12: 52  
13: 39 13: 52 14: 52 16: 39 16: 52 17: 39  
17: 52 18: 39 18: 52 19: 39 19: 52 21: 38  
22: 38 23: 39 24: 39 26: 39 27: 39 28: 39

10 29: 14 29: 39 30: 38

There are 21 hits at base# 39

There are 4 hits at base# 38

There are 3 hits at base# 31

There are 3 hits at base# 40 Ragged

15

BstNI CCwgg

30

1: 41 2: 40 5: 40 6: 40 7: 40 8: 40  
9: 40 10: 40 11: 40 12: 39 12: 53 13: 40  
13: 53 14: 53 16: 40 16: 53 17: 40 17: 53  
20 18: 40 18: 53 19: 53 21: 39 22: 39 23: 40  
24: 40 27: 40 28: 40 29: 15 29: 40 30: 39

There are 17 hits at base# 40

There are 7 hits at base# 53

There are 4 hits at base# 39

25 There are 1 hits at base# 41 Ragged

PspGI ccwgg

30

1: 41 2: 40 5: 40 6: 40 7: 40 8: 40  
9: 40 10: 40 11: 40 12: 39 12: 53 13: 40  
30 13: 53 14: 53 16: 40 16: 53 17: 40 17: 53  
18: 40 18: 53 19: 53 21: 39 22: 39 23: 40  
24: 40 27: 40 28: 40 29: 15 29: 40 30: 39

There are 17 hits at base# 40

There are 7 hits at base# 53

35 There are 4 hits at base# 39

There are 1 hits at base# 41

ScrFI CCngg

39

5 1: 41 2: 40 3: 40 3: 41 4: 40 4: 41  
5: 40 6: 32 6: 40 7: 32 7: 40 8: 40  
9: 32 9: 40 10: 40 11: 40 12: 39 12: 53  
13: 40 13: 53 14: 53 16: 40 16: 53 17: 40  
17: 53 18: 40 18: 53 19: 40 19: 53 21: 39  
22: 39 23: 40 24: 40 26: 40 27: 40 28: 40

10 29: 15 29: 40 30: 39

There are 21 hits at base# 40

There are 4 hits at base# 39

There are 3 hits at base# 41

15 MaeIII gtnac

16

1: 52 2: 52 3: 52 4: 52 5: 52 6: 52  
7: 52 9: 52 26: 52 27: 10 27: 52 28: 10  
28: 52 29: 10 29: 52 30: 52

There are 13 hits at base# 52

20

Tsp45I gtsac

15

1: 52 2: 52 3: 52 4: 52 5: 52 6: 52  
7: 52 9: 52 27: 10 27: 52 28: 10 28: 52  
29: 10 29: 52 30: 52

25 There are 12 hits at base# 52

HphI tcacc

26

30 1: 53 2: 53 3: 53 4: 53 5: 53 6: 53  
7: 53 8: 53 9: 53 10: 53 11: 59 13: 59  
14: 59 17: 59 18: 59 19: 59 20: 59 21: 59  
22: 59 23: 59 24: 59 25: 59 27: 59 28: 59  
30: 59 31: 59

There are 16 hits at base# 59

There are 10 hits at base# 53

35

BspMI ACCTGCNNNNn

14

11: 61    13: 61    14: 61    17: 61    18: 61    19: 61  
20: 61    21: 61    22: 61    23: 61    24: 61    25: 61  
30: 61    31: 61

5    There are 14 hits at base# 61    Goes into CDR1

---

11: 61    13: 61    14: 61    17: 61    18: 61    19: 61  
20: 61    21: 61    22: 61    23: 61    24: 61    25: 61  
30: 61    31: 61



Table 500: h3401-h2 captured Via CJ with BsmAI

```

! 1  2  3  4  5  6  7  8  9 10 11 12 13 14 15
! S  A  Q  D  I  Q  M  T  Q  S  P  A  T  L  S
aGT GCA Caa gac atc cag atg acc cag tct cca gcc acc ctg tct
5 ! ApaLI... a gcc acc !
L25,L6,L20,L2,L16,A11
! Extender.....Bridge...

10 ! 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
! V  S  P  G  E  R  A  T  L  S  C  R  A  S  Q
gtg tct cca ggg gaa agg gcc acc ctg tcc tgc agg gcc agt cag

15 ! 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45
! S  V  S  N  N  L  A  W  Y  Q  Q  K  P  G  Q
agt gtt agt aac aac tta gcc tgg tac cag cag aaa cct ggc cag

20 ! 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60
! V  P  R  L  L  I  Y  G  A  S  T  R  A  T  D
gtt ccc agg ctg ctg atc tat ggt gca tcc acc agg gcc act gat

25 ! 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75
! I  P  A  R  F  S  G  S  G  S  G  T  D  F  T
atc cca gcc agg ttc agt ggc agt ggg tct ggg aca gac ttc act

30 ! 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90
! L  T  I  S  R  L  E  P  E  D  F  A  V  Y  Y
ctc acc atc agc aga ctg gag cct gaa gat ttt gca gtg tat tac

35 ! 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105
! C  Q  R  Y  G  S  S  P  G  W  T  F  G  Q  G
tgt cag cgg tat ggt agc tca ccg ggg tgg acg ttc ggc caa ggg

40 ! 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120
! T  K  V  E  I  K  R  T  V  A  A  P  S  V  F
acc aag gtg gaa atc aaa cga act gtg gct gca cca tct gtc ttc

! 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135
! I  F  P  P  S  D  E  Q  L  K  S  G  T  A  S
atc ttc ccg cca tct gat gag cag ttg aaa tct gga act gcc tct

! 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150
! V  V  C  L  L  N  N  F  Y  P  R  E  A  K  V
gtt gtg tgc ctg ctg aat aac ttc tat ccc aga gag gcc aaa gta

```

```

! 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165
! Q W K V D N A L Q S G N S Q E
cag tgg aag gtg gat aac gcc ctc caa tgc ggt aac tcc cag gag

5 ! 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180
! S V T E Q D S K D S T Y S L S
agt gtc aca gag cag gac agc aag gac agc acc tac agc ctc agc

10 ! 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195
! S T L T L S K A D Y E K H K V
agc acc ctg acg ctg agc aaa gca gac tac gag aaa cac aaa gtc

! 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210
! Y A C E V T H Q G L S S P V T
15 tac gcc tgc gaa gtc acc cat cag gcc ctg agc tgc cct gtc aca

! 211 212 213 214 215 216 217 218 219 220 221 222 223
! K S F N K G E C K G E F A
20 aag agc ttc aac aaa gga gag tgt aag gcc gaa ttc gc.....

```

Table 501: h3401-d8 KAPPA captured with CJ and *BsmAI*

```

! 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
25 ! S A Q D I Q M T Q S P A T L S
aGT GCA Caa gac atc cag atg acc cag tct cct gcc acc ctg tct
! ApaLI...Extender.....a gcc acc !
L25,L6,L20,L2,L16,A11
! A GCC ACC CTG TCT ! L2

30 ! 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
! V S P G E R A T L S C R A S Q
gtg tct cca ggt gaa aga gcc acc ctc tcc tgc agg gcc agt cag
! GTG TCT CCA GGG GAA AGA GCC ACC CTC TCC TGC ! L2

35 ! 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45
! N L L S N L A W Y Q Q K P G Q
aat ctt ctc agc aac tta gcc tgg tac cag cag aaa cct gcc cag

40 ! 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60
! A P R L L I Y G A S T G A I G
gct ccc agg ctc ctc atc tat ggt gct tcc acc ggg gcc att ggt

45 ! 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75
! I P A R F S G S G S G T E F T
atc cca gcc agg ttc agt gcc agt ggg tct ggg aca gag ttc act

```

! 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90  
 ! L T I S S L Q S E D F A V Y F  
 ctc acc atc agc agc ctg cag tct gaa gat ttt gca gtg tat ttc  
 5 ! 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105  
 ! C Q Q Y G T S P P T F G G G T  
 tgt cag cag tat ggt acc tca ccg ccc act ttc ggc gga ggg acc  
 10 ! 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120  
 ! K V E I K R T V A A P S V F I  
 aag gtg gag atc aaa cga act gtg gct gca cca tct gtc ttc atc  
 15 ! 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135  
 ! F P P S D E Q L K S G T A S V  
 ttc ccg cca tct gat gag cag ttg aaa tct gga act gcc tct gtt  
 20 ! 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150  
 ! V C P L N N F Y P R E A K V Q  
 gtg tgc ccg ctg aat aac ttc tat ccc aga gag gcc aaa gta cag  
 25 ! 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165  
 ! W K V D N A L Q S G N S Q E S  
 tgg aag gtg gat aac gcc ctc caa tcg ggt aac tcc cag gag agt  
 30 ! 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180  
 ! V T E Q D N K D S T Y S L S S  
 gtc aca gag cag gac aac aag gac agc acc tac agc ctc agc agc  
 35 ! 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195  
 ! T L T L S K V D Y E K H E V Y  
 acc ctg acg ctg agc aaa gta gac tac gag aaa cac gaa gtc tac  
 40 ! 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210  
 ! A C E V T H Q G L S S P V T K  
 gcc tgc gaa gtc acc cat cag ggc ctt agc tcg ccc gtc acg aag  
 ! 211 212 213 214 215 216 217 218 219 220 221 222 223  
 ! S F N R G E C K K E F V  
 agc ttc aac agg gga gag tgt aag aaa gaa ttc gtt t

Table 508 Human heavy chains bases 88.1 to 94.2

Number of sequences..... 840

5

Number of Mismatches.....

Probe

| Id  | Ntot | 0   | 1   | 2   | 3   | 4   | 5   | 6 | 7 | Name       | Sequence.....          | Dot form.....        |
|-----|------|-----|-----|-----|-----|-----|-----|---|---|------------|------------------------|----------------------|
| 1   | 364  | 152 | 97  | 76  | 26  | 7   | 4   | 2 | 0 | VHS881-1.1 | gctgtgtattactgtgcgag   | gctgtgtattactgtgcgag |
| 2   | 265  | 150 | 60  | 33  | 13  | 5   | 4   | 0 | 0 | VHS881-1.2 | gccgtgtattactgtgcgag   | ..C.....             |
| 3   | 96   | 14  | 34  | 16  | 10  | 5   | 7   | 9 | 1 | VHS881-2.1 | gccgtatatattactgtgcgag | ..C.a.....           |
| 4   | 20   | 0   | 3   | 4   | 9   | 2   | 2   | 0 | 0 | VHS881-4.1 | gccgtgtattactgtgcgag   | ..C.....a....        |
| 5   | 95   | 25  | 36  | 18  | 11  | 2   | 2   | 0 | 1 | VHS881-9.1 | gccatgtattactgtgcgag   | ..ca.....            |
| 840 | 341  | 230 | 147 | 69  | 21  | 19  | 11  | 2 |   |            |                        |                      |
| 341 | 571  | 718 | 787 | 808 | 827 | 838 | 840 |   |   |            |                        |                      |

10

15

88 89 90 91 92 93 94 95 Codon number as in Table 195

Recognition..... Stem..... Loop.....

|              |    |                         |            |       |              |
|--------------|----|-------------------------|------------|-------|--------------|
| (VHS881-1.1) | 5' | -gctgtgtat tact-gtgcgag | cAcATccgTg | TTgTT | cAcgATgTg-3' |
| (VHS881-1.2) | 5' | -gccgtgtat tact-gtgcgag | cAcATccgTg | TTgTT | cAcgATgTg-3' |
| (VHS881-2.1) | 5' | -gccgtatat tact-gtgcgag | cAcATccgTg | TTgTT | cAcgATgTg-3' |
| (VHS881-4.1) | 5' | -gccgtgtat tact-gtgcgag | cAcATccgTg | TTgTT | cAcgATgTg-3' |
| (VHS881-9.1) | 5' | -gccatgtat tact-gtgcgag | cAcATccgTg | TTgTT | cAcgATgTg-3' |

| site of substrate cleavage

(FOKlact) 5' - cAcATccgTg TTgTT cAcgATgTg -3'

(VHEX881) 5' - AATAgTAGAc TgcAgTgTcc TcAgcccTTA AgcTgTTcAT cTgcAAGTAG -  
 AgAgTATTCT TAGAgTTgTc TcTAGAcTTA gTgAagcg-3'

! note that VHEX881 is the reverse complement of the ON below

! [RC] 5' -cgCttcacTaag-

! Scab.....

! Synthetic 3-23 as in Table 206

! |TCT|AGA|gac|aac|tot|aac|aat|act|ctc|tac|ttg|cag|atg|-

! XbaI...

! |aac|agC|TTA|AGg|gct|gag|gac|aCT|GCA|Gtc|tac|tat|t-3'

! AflII...

(VHBA881) 5' -cgCttcacTaag-

! |TCT|AGA|gac|aac|tot|aac|aat|act|ctc|tac|ttg|cag|atg|-

! |aac|agC|TTA|AGg|gct|gag|gac|aCT|GCA|Gtc|tac|tat|tgt gcg ag-3'

(VHBB881) 5' -cgCttcacTaag-

25

30

35

(VH881PCR) 5'-cgTtcaactaag|TCT|AGA|gac|aac -3'  
 |TCT|AGA|gac|aac|tct|aat|act|ctc|tac|ttg|cag|atg|-  
 |aac|agc|TTA|AGg|gct|gag|gac|aCT|GCA|Gtc|tat|tgt Acg ag-3'

5 Table 512: Kappa, bases 12-30

| ID  | Ntot | 0   | 1   | 2   | 3   | 4   | 5 | 6 | Name    | Sequence.....        | Dot Form.....       |
|-----|------|-----|-----|-----|-----|-----|---|---|---------|----------------------|---------------------|
| 1   | 84   | 40  | 21  | 20  | 1   | 2   | 0 | 0 | SK12012 | gaccagtcctccatctccc  | gaccagtcctccatctccc |
| 2   | 32   | 19  | 3   | 6   | 2   | 1   | 0 | 1 | SK12A17 | gactcagtcctccactctcc | ...t.....ct....     |
| 3   | 26   | 17  | 8   | 1   | 0   | 0   | 0 | 0 | SK12A27 | gacgcagtcctccaggcacc | ...g.....gg.a...    |
| 4   | 40   | 21  | 18  | 1   | 0   | 0   | 0 | 0 | SK12A11 | gacgcagtcctccaggcacc | ...g.....g..a...    |
| 182 | 97   | 50  | 28  | 3   | 3   | 0   | 1 |   |         |                      |                     |
| 97  | 147  | 175 | 178 | 181 | 181 | 182 |   |   |         |                      |                     |

15 URE adapters:

(SZKB1230-O12)  
 Stem..... Loop. Stem..... Recognition.....  
 5'-cACATccgTg TTgTT cACggATgTg ggAgATggAgAcTggGTC-3'  
 [RC] 5'-gaccagtcctccactctcc cACATccgTg AACAA cACggATgTg-3'  
 Recognition..... Stem..... loop. Stem.....  
 FokI.

(SZKB1230-A17)  
 Stem..... Loop. Stem..... Recognition.....  
 5'-cACATccgTg TTgTT cACggATgTg ggAgATggAgAcTggGTC-3'  
 [RC] 5'-gactcagtcctccactctcc cACATccgTg AACAA cACggATgTg-3'  
 Recognition..... Stem..... loop. Stem.....  
 FokI.

(SZKB1230-A27)  
 Stem..... Loop. Stem..... Recognition.....  
 5'-cACATccgTg TTgTT cACggATgTg ggTgCCTggAgAcTggGTC-3'  
 [RC] 5'-gacgcagtcctccaggcacc cACATccgTg AACAA cACggATgTg-3'  
 Recognition..... Stem..... loop. Stem.....  
 FokI.

(SZKB1230-A11)  
 Stem..... Loop. Stem..... Recognition.....  
 5'-cACATccgTg TTgTT cACggATgTg ggTgCCTggAgAcTggGTC-3'  
 [RC] 5'-gacgcagtcctccaggcacc cACATccgTg AACAA cACggATgTg-3'  
 Recognition..... Stem..... loop. Stem.....  
 FokI.

5' -gac cca gtc|tcc a-tc ctc c-3'  
 | Site of cleavage in substrate

What happens in the upper strand:

(SzKB1230-012\*)

5

(SzKB1230-A17\*)

(SzKB1230-A27\*)

10

(SzKB1230-A11\*)

(kapextURE)

5' -ccTctactctTgTcAcAgTgcAcAA gAc ATc cAg-3' |sense strand  
Scab.....ApalI.

15 (kapextUREPCR)

5' -ccTctactctTgTcAcAgTg-3'  
Scab.....

(kaBRO1UR)

5' -ggAggATggA cTggATgTcT TgTgcAcTgT gAcAAgAgTA gAgg-3'

[RC]

5' -ccTctactctTgTcAcAgTgcAcAA gAc ATc cAg tcc a-tc ctc c-3' ON above is R.C. of this one

20

(kaBRO2UR)

5' -ggAggATggA cTggATgTcT TgTgcAcTgT gAcAAgAgTA gAgg-3'

[RC]

5' -ccTctactctTgTcAcAgTgcAcAA gAc ATc cAg tcc a-ct ctc c-3' ON above is R.C. of this one

(kaBRO3UR)

5' -ggTgccTggA cTggATgTcT TgTgcAcTgT gAcAAgAgTA gAgg-3'

[RC]

5' -ccTctactctTgTcAcAgTgcAcAA gAc ATc cAg tcc a-gg cac c-3' ON above is R.C. of this one

(kaBRO4UR)

5' -ggTgccTggA cTggATgTcT TgTgcAcTgT gAcAAgAgTA gAgg-3'

[RC]

5' -ccTctactctTgTcAcAgTgcAcAA gAc ATc cAg tcc a-gc cac c-3' ON above is R.C. of this one  
Scab.....ApalI.

25

(VH881PCR) 5'-cgCttcacAag|TCT|AGA|gac|aac -3'  
 |TCT|AGA|gac|aac|tct|aat|act|ctc|tac|ttg|cag|atg|-  
 |aac|agC|TTA|AGg|gct|gag|gac|act|GCA|Gtc|tac|tat|tgt Acg ag-3'

5 table 512: Kappa, bases 12-30

| ID  | Ntot | 0   | 1   | 2   | 3   | 4   | 5   | 6 | Name    | Sequence.....        | Dot Form.....      |
|-----|------|-----|-----|-----|-----|-----|-----|---|---------|----------------------|--------------------|
| 1   | 84   | 40  | 21  | 20  | 1   | 2   | 0   | 0 | SK12012 | gaccagtcctccatctcc   | gaccagtcctccatctcc |
| 2   | 32   | 19  | 3   | 6   | 2   | 1   | 0   | 1 | SK12A17 | gactcagtcctccatctcc  | ...t.....ct....    |
| 3   | 26   | 17  | 8   | 1   | 0   | 0   | 0   | 0 | SK12A27 | gacgcagtcctccagccacc | ...g.....gg.a...   |
| 4   | 40   | 21  | 18  | 1   | 0   | 0   | 0   | 0 | SK12A11 | gacgcagtcctccagccacc | ...g.....g...a...  |
| 182 | 97   | 50  | 28  | 3   | 3   | 0   | 1   |   |         |                      |                    |
|     | 97   | 147 | 175 | 178 | 181 | 181 | 182 |   |         |                      |                    |

15 URE adapters:

(SZKB1230-O12) Stem..... Loop. Stem..... Recognition.....  
 5'-cAcATccgTg TTgTT cAcggATgTg ggAggATggAgAcTgggTc-3'  
 [RC] 5'-gaccagtcctccatctcc cAcATccgTg AACAA cAcggATgTg-3'  
 Recognition..... Stem..... loop. Stem.....  
 FokI.

(SZKB1230-A17) Stem..... Loop. Stem..... Recognition.....  
 5'-cAcATccgTg TTgTT cAcggATgTg ggAggATggAgAcTgAgTc-3'  
 [RC] 5'-gactcagtcctccatctcc cAcATccgTg AACAA cAcggATgTg-3'  
 Recognition..... Stem..... loop. Stem.....  
 FokI.

(SZKB1230-A27) Stem..... Loop. Stem..... Recognition.....  
 5'-cAcATccgTg TTgTT cAcggATgTg ggTgccTggAgAcTgggTc-3'  
 [RC] 5'-gacgcagtcctccagccacc cAcATccgTg AACAA cAcggATgTg-3'  
 Recognition..... Stem..... loop. Stem.....  
 FokI.

(SZKB1230-A11) Stem..... Loop. Stem..... Recognition.....  
 5'-cAcATccgTg TTgTT cAcggATgTg ggTggcTggAgAcTgggTc-3'  
 [RC] 5'-gacgcagtcctccagccacc cAcATccgTg AACAA cAcggATgTg-3'  
 Recognition..... Stem..... loop. Stem.....  
 FokI.

What happens in the upper strand:

{SZKB1230-012\*}

5'-gac cca gtc|tcc a-tc ctc c-3'

1

5

(SZKB1230-A17\*)

5'-gac tca gtc|tcc a-ct ctc c-3'

(SZKB1230-A27\*)

5'-qac qca qtc|tcc a-qq cac c-3'

IO (SZKB1230-A11\*)

5'-qac,qca qtc|tcc a-qc cac c-3'

(kapexture)

5' -ccTctactctTgTcAcAgTgcAcAA gAc ATc cAg-3' !sense strand  
Scab.....ApaI.

Scab.....ApalI:

15 (kapextUREPCR) 5' -ccTctactctTgTcAcAgTg-3'

Scab...

(kaBRO1UR)

5' -ggAggATggA cTggATgTcT TgTgcActgT gAcAAgAcTA qAqq-3'

[RC]

5' -ccTctactctTgTcAcAgTgcAcAA gAc ATc cAg tcc a-

20

(kaBRO2UR)

5' -ggAgAgTggA cTggATqTcT TqTqCAcTqT qACAAGAgTA qAqq-3'

[RC]

5' -ccTctactctTgTcAcAgTgcAcAA gAC ATc cAg tcc a-

(kaBRO3UR)

5' -ggTgccTggA cTggATgTcT TgTgcActgT gAcAAgAgTA gAgg-3'

[RC]

5' -ccTctactctTgTcAcAgTgcAcAA gAc ATc cAg tcc a-

(kaBRO4UR)

5' -ggTggcTggA cTggATgTcT TgTgcAcTgT gAcAAgAgTA gAgg-3'

cc|ctactct|g|cAcAg|gcAc



Table 515 Lambda URE adapters bases 13.3 to 19.3

| Number of mismatches..... 128              |  |    |    |    |     |     |     |     |   |
|--|--|----|----|----|-----|-----|-----|-----|---|
|  | Number of mismatches.....  |    |    |    |     |     |     |     |   |
|  | 0  | 1  | 2  | 3  | 4   | 5   | 6   | 7   | 8 |
| Id   | Ntot   |    |    |    |     |     |     |     |   |
| 1  | 58   | 45 | 7  | 1  | 0   | 0   | 2   | 2   | 1 |
| 2  | 16   | 10 | 1  | 0  | 1   | 0   | 1   | 0   | 2 |
| 3  | 17   | 6  | 0  | 0  | 4   | 1   | 1   | 5   | 0 |
| 4  | 37   | 3  | 0  | 10 | 4   | 4   | 3   | 7   | 4 |
| 128  | 64   | 8  | 11 | 5  | 8   | 5   | 11  | 11  | 5 |
| 64   | 72   | 83 | 88 | 96 | 101 | 112 | 123 | 128 |   |
| Stem..... loop. Stem..... Recognition..... |  |    |    |    |     |     |     |     |   |
| (VL133-2a2)                                | 5'-cAcATccgTg TTgTT cAcgATgTg gATcgACTgTccAggAgAc-3'               |    |    |    |     |     |     |     |   |
| [RC]                                       | 5'-gtctctggacagtcagtc <u>cAcATccgTg</u> AAcAA <u>cAcgATgTg</u> -3' |    |    |    |     |     |     |     |   |
| Recognition..... Stem..... Loop. Stem..... |  |    |    |    |     |     |     |     |   |
| Stem..... loop. Stem..... Recognition..... |  |    |    |    |     |     |     |     |   |
| (VL133-31)                                 | 5'-cAcATccgTg TTgTT cAcgATgTg gACTgTcTgTcccAaggcc-3'               |    |    |    |     |     |     |     |   |
| [RC]                                       | 5'-ggccttggacagacagtc <u>cAcATccgTg</u> AAcAA <u>cAcgATgTg</u> -3' |    |    |    |     |     |     |     |   |
| Recognition..... Stem..... Loop. Stem..... |  |    |    |    |     |     |     |     |   |
| Stem..... loop. Stem..... Recognition..... |  |    |    |    |     |     |     |     |   |
| (VL133-2c)                                 | 5'-cAcATccgTg TTgTT cAcgATgTg gACTgACTgTccAggAgAc-3'               |    |    |    |     |     |     |     |   |
| [RC]                                       | 5'-gtctctggacagtcagtc <u>cAcATccgTg</u> AAcAA <u>cAcgATgTg</u> -3' |    |    |    |     |     |     |     |   |
| Recognition..... Stem..... Loop. Stem..... |  |    |    |    |     |     |     |     |   |
| Stem..... loop. Stem..... Recognition..... |  |    |    |    |     |     |     |     |   |
| (VL133-1c)                                 | 5'-cAcATccgTg TTgTT cAcgATgTg gAcctcTgcccTggggcc-3'                |    |    |    |     |     |     |     |   |
| [RC]                                       | 5'-ggccccaggcagaggtc <u>cAcATccgTg</u> AAcAA <u>cAcgATgTg</u> -3'  |    |    |    |     |     |     |     |   |

What happens in the top strand:

```

!
! site of cleavage in the upper strand
(VL133-2a2*) 5'-g tct cct g|ga cag tcg atc
5
!
(VL133-3l*) 5'-g gcc ttg g|ga cag aca gtc
!
(VL133-2c*) 5'-g tct cct g|ga cag tca gtc
!
10 (VL133-1c*) 5'-g gcc cca g|gg cag agg gtc
!
! The following Extenders and Bridges all encode the AA sequence of 2a2 for
codons 1-15
!
! 1
15 (ON_LamEx133) 5'-ccTcTgAcTgAgT gcA cAg -
!
! 2 3 4 5 6 7 8 9 10 11 12
AGt gcT TtA acC caA ccG gcT AGT gtT AGC ggT-
!
! 13 14 15
20 tcC ccG g ! 2a2
!
! 1
(ON_LamB1-133) [RC] 5'-ccTcTgAcTgAgT gcA cAg -
!
! 2 3 4 5 6 7 8 9 10 11 12
25 AGt gcT TtA acC caA ccG gcT AGT gtT AGC ggT-
!
! 13 14 15
tcC ccG g ga cag tcg at-3' ! 2a2 N.B. the actual seq is the
30 reverse complement of the
one shown.
!
(ON_LamB2-133) [RC] 5'-ccTcTgAcTgAgT gcA cAg -
!
! 2 3 4 5 6 7 8 9 10 11 12
35 AGt gcT TtA acC caA ccG gcT AGT gtT AGC ggT-
!
! 13 14 15
tcC ccG g ga cag aca gt-3' ! 3l N.B. the actual seq is the
40 reverse complement of the
one shown.
!
(ON_LamB3-133) [RC] 5'-ccTcTgAcTgAgT gcA cAg -
!
! 2 3 4 5 6 7 8 9 10 11 12
45 AGt gcT TtA acC caA ccG gcT AGT gtT AGC ggT-
!
! 13 14 15
50 tcC ccG g ga cag tca gt -3' ! 2c N.B. the actual seq is the
reverse complement of the
one shown.
!
(ON_LamB4-133) [RC] 5'-ccTcTgAcTgAgT gcA cAg -
55

```

```

!           2   3   4   5   6   7   8   9   10  11  12
           AGt gcT TtA acC caA ccG gcT AGT gtT AGC ggT-
!
!           13  14  15
5          tcC ccG g gg cag agg gt-3' ! 1c N.B. the actual seq is the
reverse complement of the
one shown.
!
(ON_Lam133PCR) 5'-ccTcTgAcTgAgT gcA cAg AGt gc-3'

```

Table 525 ONs used in Capture of kappa light chains using CJ method and BsmAI

All ONs are written 5' to 3'.

|    |                            |  |
|----|----------------------------|--|
| 5  | REdapters (6)              | gggAggATggAgAcTgggTc                               |
|    | ON_20SK15012               | gggAAGATggAgAcTgggTc                               |
|    | ON_20SK15L12               | gggAgAgTggAgAcTgAgTc                               |
|    | ON_20SK15A17               | gggTgccTggAgAcTggcTc                               |
|    | ON_20SK15A27               | gggTggcTggAgAcTggcTc                               |
| 10 | ON_20SK15A11               | gggAgTcTggAgAcTgggTc                               |
|    | ON_20SK15B3                | gggAgTcTggAgAcTgggTc                               |
|    | Bridges (6)                | gggAggATggAgAcTgggTcATcTggATgTcTTgTgcAcTgTgAcAgAgg |
| 15 | kapbr11012                 | gggAAGATggAgAcTgggTcATcTggATgTcTTgTgcAcTgTgAcAgAgg |
|    | kapbr11L12                 | gggAgAgTggAgAcTgggTcATcTggATgTcTTgTgcAcTgTgAcAgAgg |
|    | kapbr11A17                 | gggTgccTggAgAcTgggTcATcTggATgTcTTgTgcAcTgTgAcAgAgg |
|    | kapbr11A27                 | gggTggcTggAgAcTgggTcATcTggATgTcTTgTgcAcTgTgAcAgAgg |
|    | kapbr11A11                 | gggAgTcTggAgAcTgggTcATcTggATgTcTTgTgcAcTgTgAcAgAgg |
| 20 | kapbr11B3                  | gggAgTcTggAgAcTgggTcATcTggATgTcTTgTgcAcTgTgAcAgAgg |
|    | Extender (5' biotinylated) | ccTcTgTcAcAgTgcAcAAGAcATccAgATgAcAcAgTcTcc         |
|    | kapext1bio                 |  |
|    | Primers                    |  |
| 25 | kaPCRt1                    | ccTcTgTcAcAgTgcAcAAGAc                             |
|    | kapfor                     | 5'-aca ctc tcc cct gtt gaa gct ctt-3'              |

Table 530 PCR program for amplification of kappa DNA

|      |            |
|------|------------|
| 95°C | 5 minutes  |
| 95°C | 15 seconds |
| 65°C | 30 seconds |

72°C 1 minute  
72°C 7 minutes  
4°C hold

5 Reagents (100 ul reaction):

|                      |             |
|----------------------|-------------|
| Template             | 50 ng       |
| 10x turbo PCR buffer | 1x          |
| turbo Pfu            | 4U          |
| dNTPs                | 200 µM each |
| 10 kaPCRT1           | 300 nM      |
| kapfor               | 300 nM      |

|     |             |             |             |             |            |             |
|-----|-------------|-------------|-------------|-------------|------------|-------------|
| 1   | TCCGGAGCTT  | CAGATCTGTT  | TGCCTTTTGG  | TGGGTGGTG   | CAGATCGCGT | TACGGAGATC  |
| 61  | GACCGACTGC  | TTGAGCAAAA  | GCCACGCTTA  | ACTGCTGATC  | AGGCATGGGA | TGTTATTCCG  |
| 121 | CAAAACCAATC | GTCAAGATCT  | TAAACCTGAGG | CTTTTTTTTAC | CTACTCTGCA | AGCAGCGACA  |
| 181 | TCTGGTTTGA  | CACAGAGCGA  | TCCGCGTCGT  | CAGTGGGTAG  | AAACATTAAC | ACGTTGGGAT  |
| 241 | GGCATCAATT  | TGCTTAATGA  | TCCGTGATAA  | ACTGGCAGC   | AGCCAGGCTC | TGCCATCCTG  |
| 301 | AACGTTTGGC  | TGACACGATAT | GTTGAAGGCT  | ACCGTAGTGG  | CTGCCGTACC | TATGCCATT   |
| 361 | GATAAGTTGT  | ACAGGCGCCAG | TGGCTACGAA  | ACAACCCAGG  | ACGGCCCAAC | TGGTTCGCTG  |
| 421 | AATATAAGTG  | TTGGAGCAAA  | AATTTTGAT   | GAGCGGTGC   | AGGGAGACAA | ATCACCAATC  |
| 481 | CCACAGGCGG  | TTGATCTGTT  | TGCTGGGAAA  | CCACAGCAGG  | AGGTTGTGTT | GCGTGCGCTG  |
| 541 | GAAGATACCT  | GGGAGACTCT  | TTCCAACGC   | TATGGCAATA  | ATGTGAGTAA | CTGGAANAACA |
| 601 | CCTGCAATGG  | CCTTAACGTT  | CCGGGCAAAAT | AATTTCTTTG  | GTGTACCGCA | GSCCGCAGCG  |
| 661 | GAAGAAACGC  | GTCAATCAGGC | GAGTATCAAA  | AACCGTGGAA  | CAGAAAAACA | TATGATTGTT  |
| 721 | TGTCACCAAA  | CGACAAGCGA  | TGTCCTGTG   | CTTGGCTGGG  | ATGTGCTGCG | ACCGGTCAG   |
| 781 | AGTGGGTTA   | TTGCTCCCGA  | TGGAACAGTT  | GATAAGCACT  | ATGAAGATCA | CTGAAAAATG  |
| 841 | TACGAAATT   | TTGGCCGTAA  | GTCCGCTCTGG | TTAACCAAGC  | AGGATGTGGA | GCGCATTAAG  |
| 901 | GAGTCGTCTA  | GA          |             |             |            |             |

Table 620: DNA sequence of pCES5

! pCES5 6680 bases = pCES4 with buffers in CDR1-2 and CDR3 2000.12.13

! Ngene = 6680

5 ! Useful REs (cut MAhoII fewer than 3 times) 2000.06.05

! Non-cutters

! Acc65I Ggtacc AfeI AGCgtc AvrII Cctagg  
! BsaBI GATNNnnatc BsiWI Cgtacg BsmFI Nnnnnnnnnnnnnngtccc  
10 ! BsrCI Tgtaca BstAPI GCANNNNntgc BstBI TTcgaa  
! Bst217I GTAtac Btri CACgtg Ec1136I GAGctc  
! EcoRV GARatc FseI GCCCGGcc KpnI GGTACC  
! MscI TGGcca NruI TCGgga NsiI ATGCAT  
! PacI TTAATtaa PmeI GTTaaac PmlI CACgtg  
15 ! PpuMI RGwccy PshAI GACNNNngtc SacI GAGCTc  
! SacII CCGCgg SbfI CCTGCagg SexAI Accwggg  
! SgfI GCGATcgc SnaBI TACgta SpeI Actagt  
! SphI GCATGc Sse8387I CCTGCagg StuI AGGcct  
! SwaI ATTaaat XmaI Cccggg

20 ! cutters

! Enzymes that cut more than 3 times.

! AlwNI CAGNNNctg 5  
! BglI ctgcac 4  
25 ! BstFI Rccggy 5  
! EarI CTCCTCnnnn 4  
! FnuI nNNNNNGCGGG 10

! Enzymes that cut from 1 to 3 times.

30 ! EcoO109I Rgncocy 3 7 2636 4208  
! BssSI Ctcgtg 1 12  
! "- Caccag 1 1703  
! BspHI Tcatga 3 43 148 1156  
35 ! AatII GACGTC 1 65  
! BciVI GTATCCNNNNN 2 140 1667  
! Eco57I CTGAAG 1 301  
! "- cttcag 2 1349  
! Aval Cycgrg 3 319 2347 6137  
40 ! BsiHKAI GWGCWC 3 401 2321 4245  
! HgiAI GWGCWC 3 401 2321 4245  
! BglI gcannnnnctg 1 461  
! ScaI AGTact 1 505

|    |                       |   |      |      |      |
|----|-----------------------|---|------|------|------|
| 5  | !PvuI CGATcg          | 3 | 616  | 3598 | 5926 |
|    | !FspI TGCgca          | 2 | 763  | 5946 |      |
|    | !BglI GCGNNNnggc      | 1 | 864  | 2771 | 5952 |
|    | !BpmI CTGGAG          | 3 | 898  |      |      |
|    | !-"- ctccag           | 1 | 4413 |      |      |
|    | !BsaI GGTCtCNnnn      | 1 | 916  |      |      |
|    | !AhdI GACNNNngtc      | 1 | 983  |      |      |
|    | !Eam1105I GACNNNngtc  | 1 | 983  |      |      |
| 10 | !DrdI GACNNNngtc      | 3 | 1768 | 6197 | 6579 |
|    | !SapI gaagagc         | 1 | 1998 |      |      |
|    | !PvuII CAGctg         | 3 | 2054 | 3689 | 5895 |
|    | !PflMI CCANNNntgg     | 3 | 2233 | 3943 | 3991 |
|    | !HindIII Agcgtt       | 1 | 2235 |      |      |
|    | !ApaLI Gtgcac         | 1 | 2321 |      |      |
| 15 | !BspMI Nnnnnnnngcaggt | 1 | 2328 |      |      |
|    | !-"- ACCTGCNNNNn      | 2 | 3460 |      |      |
|    | !PstI CTGCAG          | 1 | 2335 |      |      |
|    | !AccI GTmkac          | 2 | 2341 | 2611 |      |
|    | !HincII GTYrac        | 2 | 2341 | 3730 |      |
| 20 | !Sali Gtcgac          | 1 | 2341 |      |      |
|    | !TliI Ctcgag          | 1 | 2347 |      |      |
|    | !XhoI Ctcgag          | 1 | 2347 |      |      |
|    | !BbsI gtcttc          | 2 | 2383 | 4219 |      |
|    | !BlnI Gctnagc         | 1 | 2580 |      |      |
| 25 | !EspI Gctnagc         | 1 | 2580 |      |      |
|    | !SgrAI CRcgggYg       | 1 | 2648 |      |      |
|    | !AgeI Accgtg          | 2 | 2649 | 4302 |      |
|    | !AscI GGcgcgcc        | 1 | 2689 |      |      |
| 30 | !BshII Gcgcgc         | 1 | 2690 |      |      |
|    | !SfiII GGCNNNNnggcc   | 1 | 2770 |      |      |
|    | !NaeI GCCgcg          | 2 | 2776 | 6349 |      |
|    | !NcoMIV Gccggc        | 2 | 2776 | 6349 |      |
|    | !BtgI CcrYgg          | 3 | 2781 | 3553 | 5712 |
|    | !DsaI CcrYgg          | 3 | 2781 | 3553 | 5712 |
| 35 | !NcoI Coatgg          | 1 | 2781 |      |      |
|    | !StyI CcwWgg          | 3 | 2781 | 4205 | 4472 |
|    | !MfeI Caattg          | 1 | 2795 |      |      |
|    | !BspEI Tccgga         | 1 | 2861 |      |      |
|    | !BglII Agatct         | 1 | 2872 |      |      |
| 40 | !BclI Tgatac          | 1 | 2956 |      |      |
|    | !Bsu36I CCTnagg       | 3 | 3004 | 4143 | 4373 |
|    | !XcmI CCANNNNnnntgg   | 1 | 3215 |      |      |
|    | !MluI Acgcgt          | 1 | 3527 |      |      |



|    |  |   |                |
|----|--|---|----------------|
|    | HpaI GTTaaC  | 1 | 3730           |
|    | XbaI Tctaga  | 1 | 3767           |
|    |  |   |                |
| 5  | AflII Cttaag   | 1 | 3811           |
|    | BsmI NGcattc   | 1 | 3821           |
|    | - GAATGCN  | 1 | 4695           |
|    | RsrII CGgcwccg   | 1 | 3827           |
|    | NheI Gctagc  | 1 | 4166           |
|    | BstEII GgtnacC   | 1 | 4182           |
| 10 | BsmBI CTCTCNnnnn   | 2 | 4188 6625      |
|    | - Nnnnnngagacg   | 1 | 6673           |
|    | Apal GGGCCc  | 1 | 4209           |
|    | BanII GRGcYc   | 3 | 4209 4492 6319 |
|    | Bsp120I Gggccc   | 1 | 4209           |
| 15 | PspOMI Gggccc  | 1 | 4209           |
|    | BseRI NNnnnnnnnctcctc  | 1 | 4226           |
|    | - GAGGAGNNNNNNNNN  | 1 | 4957           |
|    | EcoNI CCTNNnnnagg  | 1 | 4278           |
|    | PfIFi GACNngtc   | 1 | 4308           |
| 20 | TthlIII GACNngtc   | 1 | 4308           |
|    | KasI Ggcgcc  | 2 | 4327 5967      |
|    | BstXI CCANNNNntgg  | 1 | 4415           |
|    | NotI Gggggccg  | 1 | 4507           |
|    | EagI Cggccg  | 1 | 4508           |
| 25 | BamHI Ggatcc   | 1 | 5169           |
|    | BspDI ATcgat   | 1 | 5476           |
|    | NdeI CATatg  | 1 | 5672           |
|    | EcoRI Gaattc   | 1 | 5806           |
|    | PsiI TTataa  | 1 | 6118           |
| 30 | DraIII CACNNNgTg   | 1 | 6243           |
|    | BsaAI YAcgtr   | 1 | 6246           |
|    |  |   |                |
|    | 1 gacgaagg cCTCGGata cgctatttt tataggtaa tgtcatgata ataattggtt       |   |                |
|    | BssSI.(1/2)  |   |                |
| 35 | 61 cttaGAGCTC agtggcact ttctggggaa atgtgcgag aacctctatt tgtttatttt   |   |                |
|    | AatII.   |   |                |
|    | 121 tctaaataca ttcaaatatG TATCGctca tgagacaata accctgataa atgcttcaat |   |                |
|    | BciVI..(1 of 2)  |   |                |
|    | 181 aatattgaaa aaggaagagt  |   |                |
| 40 | Base # 201 to 1061 = Apr gene from pUC19 with some RE sites removed  |   |                |
|    |  |   |                |
|    | 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15                                  |   |                |
|    | fM S I O H F R V A L I P F F A                                       |   |                |

201 atg agt att caa cat ttc cgt gtc gcc ctt att ccc ttt ttt gcg  
 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30  
 A F C L P V F A H P E T L V K  
 246 gca ttt tgc ctt cct gtt ttt gct cac cca gaa acg ctg gtg aaa  
 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45  
 V K D A E D Q L G A R V G Y I  
 291 gta aaa gat gct gaa gat cag ttg ggt gcc cga gtg ggt tac atc  
 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60  
 E L D L N S G K I L E S F R P  
 336 gaa ctg gat ctc aac agc ggt aag atc ctt gag agt ttt cgc ccc  
 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75  
 E E R F P M S T F K V L L C  
 381 gaa gaa cgt ttt cca atg atg agc act ttt aaa gtt ctg cta tgt  
 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90  
 G A V L S R I D A G Q E Q L G  
 426 ggc gcg gta tta tcc cgt att gac gcc ggg caa gaG CAa ctc ggt  
 Bcgl.....  
 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105  
 R R I H Y S Q N D L V E Y S P  
 471 Cgc cgc ata cac tat tct cag aat gac ttg gtt gAG TAC Tca cca  
 ScaI.....  
 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120  
 V T E K H L T D G M T V R E L  
 516 gtc aca gaa aag cat ctt acg gat ggc atg aca gta aga gaa tta  
 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135  
 C S A A I T M S D N T A A N L  
 561 tgc agt gct gcc ata acc atg agt gat aac act gcg gcc aac tta  
 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150  
 L L T T I G G P K E L T A F L  
 606 ctt ctg aca aCG ATC Gga gga ccg aag gag cta acc gct ttt ttg  
 PvuI.... (1/2)  
 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165  
 H N M G D H V T R L D R W E P

651 cac aac atg ggg gat cat gta act cgc ctt gat cgt tgg gaa ccg  
 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180  
 E L N E A I P N D E R D T T M  
 696 gag ctg aat gaa gcc ata cca aac gac gag cgt gac acc acg atg  
 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195  
 P V A M A T T L R K L L T G E  
 741 cct gta GCA ATG gca aca acg tTG CGC Aaa cta tta act ggc gaa  
 BsrDI...(1/2) FspI.... (1/2)  
 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210  
 L L T L A S R Q Q L I D W M E  
 786 cta ctt act cta gct tcc cgg caa caa tta ata gac tgg atg gag  
 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225  
 A D K V A G P L L R S A L P A  
 831 ggg gat aaa gtt gca gga cca ctt ctg cgc teg gcc ctt ccg gct  
 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240  
 G W F I A D K S G A G E R G S  
 876 ggc tgg ttt att gct gat aaa tCT GGA Gcc ggt gag cgt gGG TCT  
 BpmI....(1/2) BsaI....  
 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255  
 R G I I A A L G P D G K P S R  
 921 Cgc ggt atc ATT GCA gca ctg ggg cca gat ggt aag ccc tcc cgt  
 BsaI..... BsrDI...(2/2)  
 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270  
 I V V I Y T T G S Q A T M D E  
 966 atc gta gtt atc tac acG ACg ggg aGT Cag gca act atg gat gaa  
 AhdI.....  
 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285  
 R N R Q I A E I G A S L I K H  
 1011 cga aat aga cag atc gct gat ata ggt gcc tca ctg att aag cat  
 286 287  
 W .  
 1056 tgg taa  
 1062  
 1081 catatatact ttagattgat ttaaaacttc attttaatt taaaaggatc taggtgaaga  
 ctgtcagac caagtttact

```

1141      tccttttga taatcatg accaaaatcc cttaacgtga gtttcgttc cactgagcgt
1201      cagaccgcgt agaaaagatc aaagatcttt cttagatccc tttttctg cgcgtaatct
1261      gctgcttgca acaaaaaaaa ccaccgttac cagcgttggt ttgtttgcg gatcaagagc
1321      taccaactct tttccgaag gtaactggtt tcacagagc gcagatacca aatactgtcc
1381      ttctagtga gccgtagtta ggcaccact tcaagaactc ttagcaccg cctacatacc
1441      tgcgtctgct aatcctgtta ccagtggctg ctgccagtgg cgataagtcg tgtcttaccg
1501      ggttgactc aagacgatag ttaccggata aggcgcagcg gtcgggctga acggggggtt
1561      cgtgcataca gccacgcttg gagcgaacga cctacaccga actgagatac ctacagcgtg
1621      agcattgaga aagcgccacg cttcccgaa ggaagaaggc ggacagcgtat CCggttaagcg
      BciVI.. (2 of 2)
1681      gcagggtcgg aacaggagag cgCACGAGgg agcttcccagg ggaagaacgcc tggatatctt
      BssSI. (2/2)
1741      atagtcctgt cgggtttcgc cactctgac ttgagcgtgc atttttgtga tgctcgtcag
1801      gggggcgag cctatggaaa aacgccagca acgcggcctt ttacgggttc ctggcctttt
1861      gctggccttt tgctACATG Ttctttctg cgttatccc tgattctgtg gataaccgta
      PciI....
1921      ttaccgcctt tgaatgagct gataccgctc gccgcagcgg aacgaccgag cgcagcgagt
1981      cagtgagcga ggaagcgGAA GAGCgccccaa tacgcaaac gccctcccgc gcgcgttggc
      SapiI....
2041      cgattcatta atgCAGCTcg cacgacaggt ttccgactg gaaagcgggc agtgagcgca
      PvuII. (1/3)
2101      acgcaatTAA TGTgagtttag ctcactcatt aggcacccca ggcTTTACac ttatgcttc
      ..-35..   Plac
2161      cggctcgtat gttgtgtgga attgtgagcg gataacaatt tcacaCAGGA AACAGCTATG
      M13Rev_seq_primer
2221      Accatgatta cgCCAAGCTT TGGagccttt tttttggaga ttttcaac
      PflMI.....
      Hind3.
signal::linker::CLight
30      1  2  3  4  5  6  7  8  9 10 11 12 13 14 15
      fm K K L L L F A I P L V V P F Y
      gtg aaa aaa tta tta ttc gca att cct tta gtt gtt cct ttc tat

35      16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
      S H S A Q V Q L Q V D L E I K
      tct cac aGT GCA Cag gtc caa CTG CAG GTC GAC CTC GAG atc aaa
      ApaLI..... PstI... XhoI...
      BspMI...
      Sali...
      AccI... (1/2)
      HincII. (1/2)
40

```

Vlight domains could be cloned in as ApaLI-XhoI fragments.  
VL-CL( $\kappa$ ) segments can be cloned in as ApaLI-AscI fragments. <-----

```

2701      ctatttcaag gagacagtca ta
!
!
!
5      PeIB::3-23(stuffed)::CH1::III fusion gene
!
!
!
1      1  2  3  4  5  6  7  8  9  10 11 12 13 14 15
!
!
!
2723      M K Y L L P T A A A G L L L L
!
!
!
2723      atg aaa tac.cta ttg cct acg gca gcc gct gga ttg tta tta ctc
!
!
!
10      -----
!
!
!
16      16 17 18 19 20 21 22
!
!
!
2768      A A Q P A M A
!
!
!
2768      gcG GCC cag ccG GCC atg gcc
!
!
!
15      SfiI.....
!
!
!
2768      NgoMIV..(1/2)
!
!
!
2768      NcoI.....
!
!
!
20      FR1(DP47/V3-23)-----
!
!
!
2789      23 24 25 26 27 28 29 30
!
!
!
2789      E V Q L L E S G
!
!
!
2789      gaa|gtt|CAA|TTG|tta|gag|tct|ggt|
!
!
!
2789      | MfeI |
!
!
!
25      -----FR1-----
!
!
!
2813      31 32 33 34 35 36 37 38 39 40 41 42 43 44 45
!
!
!
2813      G G L V Q P G G S L R L S C A
!
!
!
2813      |ggc|ggg|ctt|gtt|cag|cct|ggg|ggg|tct|tta|cgt|ctt|tct|tgc|gct|
!
!
!
30      -----FR1-----
!
!
!
2858      46 47 48
!
!
!
2858      A S G
!
!
!
2858      |gct|TCC|GGA|
!
!
!
2858      | BspEI |
!
!
!
35      Stuffer for CDR1, FR2, and CDR2----->
!
!
!
2867      There are no stop codons in this stuffer.
!
!
!
2867      gcttcAGATC Tggttgccct
!
!
!
2867      BglII..
!
!
!
2887      tttgtggggt ggtgcagatc gcgttacgga gatcgaccga ctgcttgagc aaagccacg
!
!
!
2947      cttaactgct GATCAGgcgat gggatgttat tcgccaaacc agtcgtcagg atcttaacct
!
!
!
2947      Bell...
!
!
!
3007      gaggcgtttt ttacctactc tgcaagcagc gacatctggt ttgacacaga gcgatccgcg
!
!
!
3067      tcgtcaggtg gtagaacat taacacgttg ggaatggcat aatttgctta atgatgatgg
!
!
!
40

```

1000 1000 1000 1000 1000 1000 1000 1000 1000 1000  
 1000 1000 1000 1000 1000 1000 1000 1000 1000 1000  
 1000 1000 1000 1000 1000 1000 1000 1000 1000 1000

3127 taaaaacctgg cagcagccag gctctgccat cctgaacgtt tgctgacca gtatgttgaa  
 3187 ggtacccgta gtggtgcgg tacctatgcc Atttgataag TGGtacagcg ccagtggcta  
 3247 cgaacaacc caggacggcc caactggttc gctgaatata agtgttgagg caaaaatttt  
 3307 gtatgagcg gtgcaggagg acaaatcacc aatccacag gcggttgatc tgtttgctgg  
 3367 gaaaccacag caggaggttg tgttgctgc gctggaagat acctgggaga ctctttccaa  
 3427 acgctatggo aataatgtga gtaactgtaa aacacctgca atggccttaa cgttcggggc  
 3487 aaataatttc tttgtgtac cgcagccgc agcgaagaa ACCGTCatc aggcggagta  
 3547 tcaaaacctg ggaacagaaa acgatatgat tgtttctca ccaacgaaa gcgacgtcc  
 3607 tgtgctgcc tgggatgtgg tcgcaccgg tcagagtggg tttattgctc ccgatggaac  
 3667 agttgataag cactatgaag atcagctgaa aatgracgaa aatttgcc gtaagtcgt  
 3727 ctgGTTAACg aagcaggatg tggaggcgca taaggagtgc  
 HpaI..  
 HincII(2/2)

-----FR3-----  
 4 5 6 7 8 9 10 11 12 13 14 15 16  
 93 94 95 96 97 98 99 100 101 102 103 104 105  
 S R D N S K N T L Y L Q M  
 |TCT|AGA|gac|aac|tct|aag|aat|act|ctc|tac|ctg|cag|atg|  
 | XbaI |

---PR3----->  
 17 18 19 20  
 106 107 108 109  
 N S L S l s i r s g  
 |aac|agC|TTA|AG t ctg agc att CGG TCC G  
 |AflII | RsrII..

q h s p t  
 3834 gg caa cat tct cca aac tga ccagacga cacaaacggc  
 3872 ttacgtataa tccgcgcgat gggatgtaa agagtgggc tctttgctgg cctgactca  
 3932 tcagatgaag gccaaaaatt ggcaggagtg gacacagcag gcagcgaac aagcactgac  
 3992 catcaactgg tactatgctg atgtaaacgg caatatgggt tatgttcata ctggtgctta  
 4052 tccagatcgt caatcagccc atgacccgg attaccgggt cctggtacgg gaaaatggga  
 4112 ctggaaaggg ctattgcctt ttgaaatgaa ccctaagggt tataacccc ag  
 4164 aa GCTAGC ctgcggcttc  
 NheI..

4182 G|GTC|ACC| gtc tca agc  
 | BstEII |

```

5      136 137 138 139 140 141 142 143 144 145 146 147 148 149 150
      A S T K G P S V F P L A P S S
4198 gcc tcc acc aag ggc cca tgc ttc ccc ctg gca ccc tcc tcc

      151 152 153 154 155 156 157 158 159 160 161 162 163 164 165
      K S T S G G T A A L G C L V K
4243 aag agc acc tct ggg ggc aca gcg gcc ctg ggc tgc ctg gtc aag

      166 167 168 169 170 171 172 173 174 175 176 177 178 179 180
      D Y F P E P V T V S W N S G A
4288 gac tac ttc ccc gaa ccg gtg acg gtg tgc tgg aac tca ggc gcc

      181 182 183 184 185 186 187 188 189 190 191 192 193 194 195
      L T S G V H T F P A V L Q S S
4333 ctg acc agc ggc gtc cac acc ttc ccg gct gtc cta cag tcc tca

      196 197 198 199 200 201 202 203 204 205 206 207 208 209 210
      G L Y S L S S V V T V P S S S
4378 gga ctc tac tcc ctc agc agc gta gtg acc gtg ccc tcc agc agc

      211 212 213 214 215 216 217 218 219 220 221 222 223 224 225
      L G T Q T Y I C N V N H K P S
4423 ttg ggc acc cag acc tac atc tgc aac gtg aat cac aag ccc agc

      226 227 228 229 230 231 232 233 234 235 236 237 238
      N T K V D K K V E P K S C
4468 aac acc aag gtg gac aag AAA GTT GAG CCC AAA TCT TGT
      ON-TQHCforw.....
      Poly His linker
      139 140 141 142 143 144 145 146 147 148 149 150
      A A A H H H H H G A A
      GCG GCC GCA cat cat cac cat cac ggc gcc gca
      NotI.....
      EagI.....

35      151 152 153 154 155 156 157 158 159 160 161 162 163 164 165
      E Q K L I S E E D L N G A A
4543 gaa caa aaa ctc atc tca gaa gag gat ctg aat ggc gcc tag

      Mature III----->....
      166 167 168 169 170 171 172 173 174 175 176 177 178 179 180

```



4588 T V E S C L A K P H T E N S F  
 act gtt gaa agt tgt tta gca aaa cct cat aca gaa aat tca ttt  
 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195  
 T N V W K D D K T L D R Y A N  
 4633 act aac gtc tgg aaa gac gac aaa act tta gat cgt tac gct aac  
 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210  
 Y E G C L W N A T G V V C T  
 4678 tat gag ggc tgt ctg tgg AAT GCT aca ggc gtt gtg gtt tgt act  
 BsmI....  
 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225  
 G D E T Q C Y G T W V P I G L  
 4723 ggt gac gaa act cag tgt tac ggt aca tgg gtt cct att ggg ctt  
 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240  
 A I P E N E G G G S E G G S  
 4768 gct atc cct gaa aat gag ggt ggt ggc tct gag ggt ggc ggt tct  
 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255  
 E G G G S E G G G T K P P E Y  
 4813 gag ggt ggc ggt tct gag ggt ggc ggt act aaa cct cct gag tac  
 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270  
 G D T P I P G Y T Y I N P L D  
 4858 ggt gat aca cct att ccg ggc tat act tat atc aac cct ctc gac  
 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285  
 G T Y P P G T E Q N P A N P N  
 4903 ggc act tat ccg cct ggt act gag caa aac ccc gct aat cct aat  
 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300  
 P S L E E S Q P L N T F M F Q  
 4948 cct tct ctt GAG GAG tct cag cct ctt aat act ttc atg ttt cag  
 BseRI... (2/2)  
 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315  
 N N R F R N R Q G A L T V Y T  
 4993 aat aat agg ttc cga aat agg cag ggt gca tta act gtt tat acg  
 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330  
 G T V T Q G T D P V K T Y Y Q

5038 ggc act gtt act caa ggc act gac ccc gtt aaa act tat tac sag  
 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345  
 Y T P V S S K A M Y D A Y W N  
 5083 tac act cct gta tca tca aaa gcc atg tat gac gct tac tgg aac  
 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360  
 G K F R D C A F H S G F N E D  
 5128 ggt aaa ttc aga gac tgc gct ttc cat tct ggc ttt aat gag CAT  
 BanHI...  
 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375  
 P F V C E Y Q G Q S S D L P Q  
 5173 Cca ttc gtt tgt gaa tat caa ggc caa tgc tct gac CTG Cct caa  
 BspMI...(2/2)  
 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390  
 P P V N A G G G S G G G S G G  
 5218 cct cct gtc aat gct ggc ggc ggc tct ggt ggt ggt tct ggt ggc  
 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405  
 G S E G G G S E G G G S E G G  
 5263 ggc tct gag ggt ggc ggc tct gag ggt ggc tct gag ggt ggc  
 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420  
 G S E G G G S G G G S G S G D  
 5308 ggc tct gag ggt ggc ggt tcc ggt ggc ggc tcc ggt tcc ggt gat  
 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435  
 F D Y E K M A N A N K G A M T  
 5353 ttt gat tat gaa aaa atg gca aac gct aat aag ggg gct atg acc  
 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450  
 E N A D E N A L Q S D A K G K  
 5398 gaa aat gcc gat gaa aac gcg cta cag tct gac gct aaa ggc aaa  
 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465  
 L D S V A T D Y G A A I D G F  
 5443 ctt gat tct gtc gct act gat tac ggt gct gct ATC GAT ggt ttc  
 BspDI...  
 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480  
 I G D V S G L A N G N G A T G

5488 att ggt gac gtt tcc ggc ctt gct aat ggt aat ggt gct act ggt  
 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495  
 D F A G S N S Q M A Q V G D G  
 5533 gat ttt gct ggc tct aat tcc caa atg gct caa gtc ggt gac ggt  
 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510  
 D N S P L M N N F R Q Y L P S  
 5578 gat aat tca cct tta atg aat aat ttc cgt caa tat tta cct tct  
 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525  
 L P Q S V E C R P Y V F G A G  
 5623 ttg cct cag tcg gtt gaa tgt cgc cct tat gtc ttt ggc gct ggt  
 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540  
 K P Y E F S I D C D K I N L F  
 5668 aaa cCA TAT Gaa ttt tct att gat tgt gac aaa ata aac tta ttc  
 NdeI....  
 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555  
 R G V F A F L L Y V A T F M Y  
 5713 cgt ggt gtc ttt gcg ttt ctt tta tat gtt gcc acc ttt atg tat  
 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570  
 V F S T F A N I L R N K E S  
 5758 gta ttt tcg acg ttt gct aac ata ctg cgt aat aag gag tct taa  
 571  
 5803 taa GAATTC  
 EcoRI.  
 5812 actggcgt cgttttaca cgtcgtgact gggaaaacc tggcgttacc caacttaac  
 5871 gccttgagc acatccccc ttgcgcagct ggcgtaatag cgaagaggcc gcacCGATC  
 PvuII..  
 5931 Gcccttcca acagtTGGC Agcctgaatg gcgaatGGC Cctgatggg tattttctcc  
 ...PvuII... (3/3) FspI... (2/2) KsiI... (2/2)  
 5991 ttacgatct gtgcgtatt tcacaccgca tataaattgt aaacgttaat attttgttaa  
 6051 aattcggtt aaattttgt taatcagct catttttaa ccaataggcc gaaatcgga  
 6111 aaatccCTA TAAatcaaaa gaatgccc agatagggtt gagtgttgtt ccagtttga  
 PsiI....  
 6171 acaagagtc actattaaag aacgtgact ccaactcaa aggcgaaaa accgtctatc  
 6231 agggcgatgg ccCACTacGT Gaaccatcac ccaatcaag tttttggg tcgagtgcc  
 DraIII....

6291 gtaaaagcact aaatcggaac cctaaaggga gcccccgatt tagagcttga cggggaaaaGC  
 6351 CGCGgaacgt ggcgagaaag gaagggaaga aagcgaagg agcgggcgt agggcgctgg  
 6411 caagttagc ggtcagctg cgcgtaacca ccacaccgc cgcgcttaac gcgcgcctac  
 6471 agggcgctga ctatggtgc tttagcgggt gcagtctcag tacaatctgc totgatgccg  
 6531 catagttaag ccagcccgga caccgcca caccgctga cgcgccctga cgggcttgc  
 6591 tgctccggc atccgcttac agacaagctg tgaccgtctc cgggagctgc atgtgtcaga  
 6651 ggtttccacc gtcacaccg aaacgcgcga

6291 gtaaaagcact aaatcggaac cctaaaggga gcccccgatt tagagcttga cggggaaaaGC  
 6351 CGCGgaacgt ggcgagaaag gaagggaaga aagcgaagg agcgggcgt agggcgctgg  
 6411 caagttagc ggtcagctg cgcgtaacca ccacaccgc cgcgcttaac gcgcgcctac  
 6471 agggcgctga ctatggtgc tttagcgggt gcagtctcag tacaatctgc totgatgccg  
 6531 catagttaag ccagcccgga caccgcca caccgctga cgcgccctga cgggcttgc  
 6591 tgctccggc atccgcttac agacaagctg tgaccgtctc cgggagctgc atgtgtcaga  
 6651 ggtttccacc gtcacaccg aaacgcgcga

5

Table 630: Oligonucleotides used to clone CDR1/2 diversity

All sequences are 5' to 3'.

5 1) ON\_CD1Bsp, 30 bases

A C C T C A C T g g c T T C C g g A  
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18

10 T T C A C T T T C T C T  
19 20 21 22 23 24 25 26 27 28 29 30

2) ON\_Br12, 42 bases

15 A g A A A C C C A C T C C A A A C C  
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18

T T T A C C A g g A g C T T g g C g  
19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36

A A C C C A  
37 38 39 40 41 42

3) ON\_CD2Xba, 51 bases

25 g g A A A g g C A g T g A T C T A g A  
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18

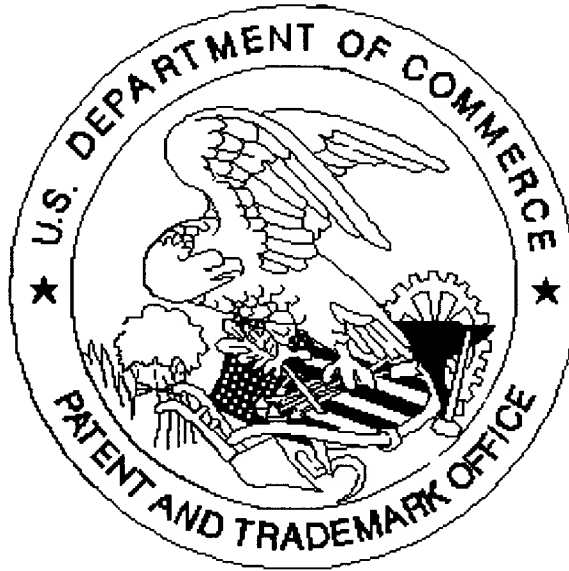
g A T A g T g A A g C g A C C T T T  
30 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36

A A C g g A g T C A g C A T A  
37 38 39 40 41 42 43 44 45 46 47 48 49 50 51

35 4) ON\_BotXba, 23 bases



United States Patent & Trademark Office  
Office of Initial Patent Examination -- Scanning Division



SCANNED # 8

Application deficiencies found during scanning:

☐ Page(s) \_\_\_\_\_ of \_\_\_\_\_ were not present  
for scanning. (Document title)

☐ Page(s) \_\_\_\_\_ of \_\_\_\_\_ were not present  
for scanning. (Document title)

*Specification starts from page 2.*

☐ ***Scanned copy is best available.***